Domain, Convexity

Computational Intelligence, Lecture 4

by Sergei Savin

Spring 2024

Content

- Problems with inequality constraints
- Quadratic programming
- Domain. Bounded and unbounded domains.
- Convex domains
- Convex combination
- Supporting hyperplane
- Convex functions
- Jensen's inequality
- Epigraph

Problems with inequality constraints

Problem 1.

$$\begin{aligned} \underset{\mathbf{x}}{\text{minimize}} \quad & ||\mathbf{D}\mathbf{x} + \mathbf{f}||, \\ \text{subject to} \quad & \mathbf{x} \leq \mathbf{b}. \end{aligned} \tag{1}$$

Problem 2.

$$\begin{aligned} \underset{\mathbf{x}}{\text{minimize}} \quad & ||\mathbf{x}||, \\ \text{subject to} \quad & \mathbf{A}\mathbf{x} \leq \mathbf{b}. \end{aligned} \tag{2}$$

Problem 3.

$$\begin{aligned} \underset{\mathbf{x}}{\text{minimize}} \quad & ||\mathbf{D}\mathbf{x} + \mathbf{f}||, \\ \text{subject to} \quad & \begin{cases} \mathbf{A}\mathbf{x} \leq \mathbf{b}, \\ \mathbf{C}\mathbf{x} = \mathbf{d}. \end{cases} \end{aligned} \tag{3}$$

Mentioned problems can be described together as quadratic programs. The name is due to the cost function being quadratic (or equivalent). They are allowed to have linear equality and inequality constraints.

General form of a quadratic program is given below:

$$\begin{aligned} \underset{\mathbf{x}}{\text{minimize}} \quad & \mathbf{x}^\top \mathbf{H} \mathbf{x} + \mathbf{f}^\top \mathbf{x}, \\ \text{subject to} \quad & \begin{cases} \mathbf{A}\mathbf{x} \leq \mathbf{b}, \\ \mathbf{C}\mathbf{x} = \mathbf{d}. \end{cases} \end{aligned} \tag{4}$$

where $\mathbf{H}$ is positive-definite and $\mathbf{A}\mathbf{x} \leq \mathbf{b}$ describe a *convex region*.

Domain, Convexity

Problem 1. Find minimum of the function $f = x^2 + 2y^2$ if $x \in \mathbb{R}$ and $y \in \mathbb{R}$.

Problem 2. Find minimum of the function $f = x^2 + 2y^2$ if $x \in [1\ 2]$ and $y \in [2\ 5]$.

Note that solutions of problems 1 and 2 are different, and this is only due to the difference of the allowed values that the *decision variables* x and y can assume.

Definition 1

Space of all allowed values that decision variables can assume is called the *domain* of optimization problem.

Problem 3. Find minimum of the function $f = -x^2$ if $x \in [-3\ 2]$.

Problem 4. Find minimum of the function $f = -x^2$ if $x \in \mathbb{R}$.

Problem 5. Find minimum of the function $f = -x^2$ if $x \in (-\infty\ 2]$.

The major difference between domains of the problems 2, 3 vs problems 1, 4 and 5 is that the latter are *not bound* (i.e., you can construct a sequence of the values in the domain that would approach infinity).

We can see that in the case of problems 3-5, bounded domain allows the problem to have a solution.

Problem 6. Find maximum of the function $f = x^2$ if $1 \leq x < 2$.

Problem 7. Find minimum of the function $f = x^2$ if $1 \leq x < 2$.

This time, it is the fact that one of the *boundaries* of the domain was not included in the domain that has lead the problem 6 to have no solution, while problem 7 had one. For the problem 6 we can pick a value arbitrary close to $x = 2$, approaching it from the left, but for any such value, there always will be other values of the decision variable closer to $x = 2$ and hence producing larger values of f.

Definition 2

Domain is *convex* iff for any two points in the domain, the line segment connecting them is also in the domain.

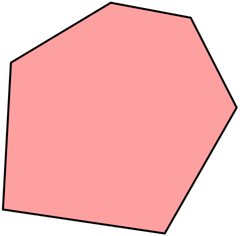

Figure 1: Convex domain

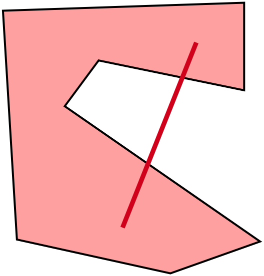

Figure 2: Non-convex domain

Convex combination, 1

In the proofs it is convenient to remember that for any two points $\mathbf{x}_1$ and $\mathbf{x}_2$, all points in the line segment connecting them are given as $\mathbf{x}_l = \alpha \mathbf{x}_1 + (1 - \alpha)\mathbf{x}_2$, where $\alpha \in [0\ 1]$. This is called *convex combination.*

We can think of the variable α as a slider - sliding α from 0 to 1 we slide $\mathbf{x}_l$ from $\mathbf{x}_2$ to $\mathbf{x}_1$.

A line segment between points $\mathbf{x}_1$ and $\mathbf{x}_2$ is called *convex hull* of points $\mathbf{x}_1$ and $\mathbf{x}_2$. It can be defined as:

$$\mathcal{L} = \{\mathbf{x} : \ \alpha \mathbf{x}_1 + (1 - \alpha)\mathbf{x}_2, \ \alpha \in [0\ 1]\} \tag{5}$$

Equivalently, a convex combination (line segment) of two points $\mathbf{x}_1$ and $\mathbf{x}_2$ is given as:

$$\mathcal{L} = \{\mathbf{x} : \ \beta_1 \mathbf{x}_1 + \beta_2 \mathbf{x}_2, \ \beta_1 + \beta_2 = 1, \ \beta_{1,2} \in [0 \ 1]\} \tag{6}$$

To show the equivalence, we can observe that $\beta_1 = 1 - \beta_2$. Defining $\alpha = \beta_1$ we get $\beta_2 = 1 - \alpha$, with $\alpha \in [0 \ 1]$ demonstrating equivalence.

$\mathbf{x} \in \mathcal{X}$, $\mathcal{X} = \mathbb{R}^n$ is convex.

$\mathbf{x} \in \mathcal{X}$, $\mathcal{X} = \{\mathbf{x} \in \mathbb{R}^n : \ \mathbf{x} \leq \mathbf{h}\}$ is convex.

Proof: Note that $\alpha \mathbf{x}_1 \leq \alpha \mathbf{h}$ and $(1-\alpha)\mathbf{x}_2 \leq (1-\alpha)\mathbf{h}$, hence, $\alpha \mathbf{x}_1 + (1-\alpha)\mathbf{x}_2 \leq \alpha \mathbf{h} + (1-\alpha)\mathbf{h} = \mathbf{h}$. $\square$

$\mathbf{x} \in \mathcal{X}$, $\mathcal{X} = \{\mathbf{x} \in \mathbb{R}^2 : \ x_1^2 + x_2^2 \leq h^2\}$ is convex.

Proof: This is the same as $||\mathbf{x}|| \leq h$. Note that $||\alpha \mathbf{x}_1|| \leq \alpha h$ and $||(1-\alpha)\mathbf{x}_2|| \leq (1-\alpha)h$. Applying triangle inequality, we get:

$$||\alpha \mathbf{x}_1 + (1-\alpha)\mathbf{x}_2|| \leq ||\alpha \mathbf{x}_1|| + ||(1-\alpha)\mathbf{x}_2|| \leq$$
$$\leq \alpha h + (1-\alpha)h = h$$

So the convex combination of $\mathbf{x}_1$ and $\mathbf{x}_2$ is still in the domain. $\square$

$\mathbf{x} \in \mathcal{X}$, $\mathcal{X} = \{\mathbf{x} \in \mathbb{R}^n : \ \mathbf{x}^\top \mathbf{H} \mathbf{x} \leq 1\}$, where $\mathbf{H} \succ 0$ is positive-definite symmetric matrix is convex.

For any positive-definite symmetric $\mathbf{H}$ it is true that $\mathbf{H} = \mathbf{D}^\top \mathbf{D}$, where $\mathbf{D} = \sqrt{\mathbf{H}}$ is called a matrix square root and it is full rank. With that $\mathbf{x}^\top \mathbf{H} \mathbf{x} \leq 1$ becomes $\mathbf{x}^\top \mathbf{D}^\top \mathbf{D} \mathbf{x} \leq 1$. Defining $\mathbf{y} = \mathbf{D}\mathbf{x}$ we get $\mathcal{X} = \{\mathbf{D}^{-1}\mathbf{y} : \ \mathbf{y}^\top \mathbf{y} \leq 1\}$. This is a linearly deformed previously covered domain, and as such it is also convex.

Examples of non-convex domains

$x \in \mathcal{X}$, $\mathcal{X} = [-1 \ 2] \cup [3 \ 7]$ is not convex..

$\mathbf{x} \in \mathcal{X}$, $\mathcal{X} = \{\mathbf{x} \in \mathbb{R}^2 : \ x_1^2 + x_2^2 \geq h^2\}$ is not convex. Prove it.

$\mathbf{x} \in \mathcal{X}$, $\mathcal{X} = \{\mathbf{x} \in \mathbb{R}^n : \ \mathbf{x}^\top \mathbf{H} \mathbf{x} \geq 1\}$, where $\mathbf{H}$ is positive-definite symmetric matrix is not convex. Prove it.

These proves simply require one counter-example to show that the defining property of convex domains does not hold.

For a convex set $\mathcal{C}$ and a point $\mathbf{x}_0$ on its boundary, if inequality $\mathbf{a}^\top \mathbf{x} \leq \mathbf{a}^\top \mathbf{x}_0$ holds for $\forall x \in \mathcal{C}$, then $\mathbf{a}^\top \mathbf{x} = \mathbf{a}^\top \mathbf{x}_0$ is a *supporting hyperplane*.

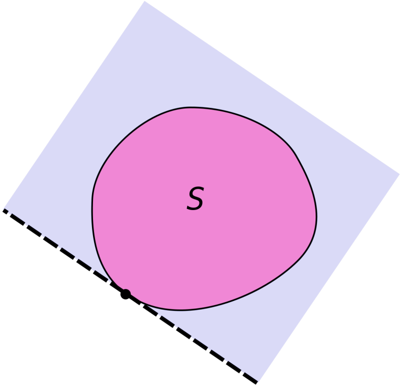

Figure 3: Supporting hyperplane (source: wikipedia)

Convex functions

Definition 3

Function $f(\mathbf{x})$ defined on a domain $\mathcal{D}$, for which it holds that $\forall \mathbf{x}_1, \mathbf{x}_2 \in \mathcal{D}$, $f(\alpha \mathbf{x}_1 + (1-\alpha)\mathbf{x}_2) \leq \alpha f(\mathbf{x}_1) + (1-\alpha) f(\mathbf{x}_2)$ is called *a convex function.*

Figure 4: Convex function

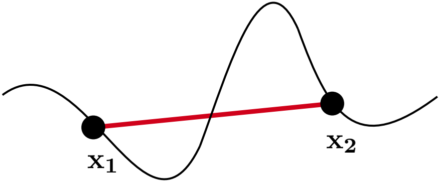

Figure 5: Non-convex function

The inequality we studied:

$$f(\alpha \mathbf{x}_1 + (1-\alpha)\mathbf{x}_2) \leq \alpha f(\mathbf{x}_1) + (1-\alpha) f(\mathbf{x}_2) \tag{7}$$

is called *Jensen's inequality*. It can be re-written equivalently as:

$$\begin{cases} f(\alpha_1 \mathbf{x}_1 + \alpha_2 \mathbf{x}_2) \leq \alpha_1 f(\mathbf{x}_1) + \alpha_2 f(\mathbf{x}_2) \\ \alpha_1 + \alpha_2 = 1 \end{cases} \tag{8}$$

Epigraph of a function $\varphi(x)$ is a set of points "above the graph":

$$\text{epi}(\varphi) = \{(x, t) : \ \varphi(x) \leq t\} \tag{9}$$

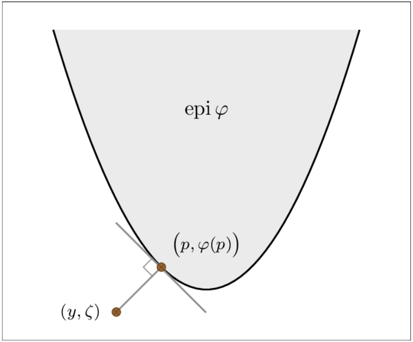

Figure 6: Epigraph. Credit - link.

Epigraph of a convex function is convex. A tangent to a function is a supporting hyperplane to its epigraph.

Here are some single-variable convex functions:

- $f(x) = 1$
- $f(x) = x$; $f(x) = x + 1$, $f(x) = 6x + 3$
- $f(x) = x^2$; $f(x) = (x - 5)^2$; $f(x) = (x + 1)^2 - 10$
- $f(x) = x^3$, if $x > 0$

Here are some multi-variable convex functions:

- $f(\mathbf{x}) = \mathbf{A}\mathbf{x} + \mathbf{b}$
- $f(\mathbf{x}) = \mathbf{x}^\top \mathbf{H}\mathbf{x}$, $\mathbf{H} \succ 0$
- $f(\mathbf{x}) = \max(x_1,\ ...\ x_n)$.
- $f(\mathbf{x}) = \log(e^{x_1} + ... + e^{x_n})$.
- $f(x, y) = \frac{x^2}{y}$.

Definition 3

If the domain of the optimization problem is convex and the cost function is convex, it is called a *convex optimization problem.*

Additionally, we will always assume that the domain of the convex optimization problem contains its boundary. Also, without the loss of generality, we will consider only minimization problems.

There are a few important properties of convex optimization problems (with our additional assumption):

- If the domain is non-empty, there is a solution.
- The problem has no local minima. We can find a path from any point to the solution, along which the cost function will not increase.

Lecture slides are available via Github, links are on Moodle:

github.com/SergeiSa/Computational-Intelligence-2024