Harvard CS 182: Artificial Intelligence September 28, 2022

Section 4: Convex Optimization, Integer Programming

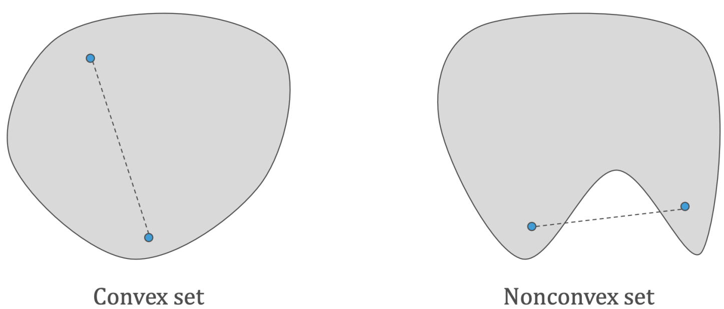
Lecturer: Ariel Procaccia *Author: Max Guo, Eric Helmold, Christopher Lee*

1 Convex Optimization

Definition 1 (Convex Optimization Problem) *A convex optimization problem is a specialization of a general optimization problem* $\min_{\mathbf{x}} f(\mathbf{x})$ *such that* $\mathbf{x} \in \mathcal{F}$ *where the objective function* $f : \mathbb{R}^n \to \mathbb{R}$ *is a convex function, and the feasible region* $\mathcal{F}$ *is a convex set.*

1.1 Convex Sets

Definition 2 (Convex Sets) *A set* $\mathcal{F} \subseteq \mathbb{R}^n$ *is convex if for all* $\mathbf{x}, \mathbf{y} \in \mathcal{F}$ *and* $\theta \in [0, 1]$, $\theta\mathbf{x} + (1-\theta)\mathbf{y} \in \mathcal{F}$. *In other words, a set is convex if and only if all convex combinations of any two elements results in another element contained within the set.*



Intuitively the definition of convex sets is: for any two points in the set, if I draw the line between the two points, every point in the line must be in the set. Now we'll provide some examples of convex sets:

Example 1 *Prove that*

$$\mathcal{F} = \{\mathbf{x} \in \mathbb{R}^n : \forall i = 1, \ldots, n, a \leq x_i \leq b\}$$

is a convex set.

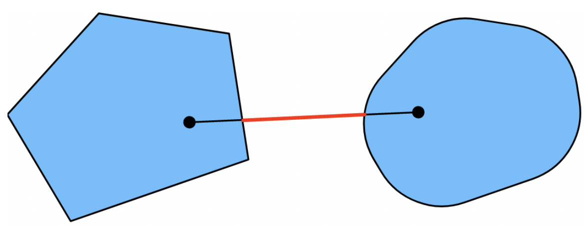
Proof: Let $\mathbf{x}, \mathbf{y} \in \mathcal{F}$, and $\theta \in [0,1]$. For all $i = 1, \ldots, n$, $a \leq x_i$ and $a \leq y_i$, so $\theta x_i + (1-\theta)y_i \geq \theta a + (1-\theta)a = a$. Similarly, $\theta x_i + (1-\theta)y_i \leq \theta(b) + (1-\theta)b = b$. Therefore, $\theta \mathbf{x} + (1-\theta)\mathbf{y} \in \mathcal{F}$. ■

Moreover, *intersections of convex sets are convex.* An argument works likes this: for any two points within the intersection of several convex sets, the line between them is (because each of the sets is convex) also withing each of the convex sets, so the line must be in the intersection, which proves the intersection is a convex set.

Note that it is *not* true that the union of convex sets is convex.

Problem 1 *Give an example where the union of convex sets is not convex.*

1. Consider two convex sets S_1 and S_2 that do not intersect. Their union will not be convex because the line segment connecting a point in S_1 to a point in S_2 will traverse through $(S_1 \cup S_2)^C$, and thus $S_1 \cup S_2$ is not convex.



Problem 2 *Show that a set is convex if and only if its intersection with any line is convex.*

2. Suppose that set S is convex and choose any line ℓ. If $S \cap \ell = \emptyset$, then the forward condition holds trivially since the empty set is convex, so assume the intersection is not empty. Choose any two points $\mathbf{x}, \mathbf{y} \in S \cap \ell$ and $\theta \in [0,1]$. By the definition of convexity, $\theta \mathbf{x} + (1-\theta)\mathbf{y} \in S$. Furthermore, by definition of a line, $\theta \mathbf{x} + (1-\theta)\mathbf{y} \in \ell$. Thus,

$$\theta \mathbf{x} + (1-\theta)\mathbf{y} \in S \cap \ell,$$

and thus $S \cap \ell$ is convex.

For the converse, suppose that S is some set, not necessarily convex. Suppose we take any two arbitrary points $\mathbf{x}, \mathbf{y} \in S$ with $\mathbf{x} \neq \mathbf{y}$, and let ℓ be the line that connects $\mathbf{x}$ with $\mathbf{y}$. By the statement of the converse, we have that $S \cap \ell$ is convex. This means that for any $\theta \in [0,1]$, $\theta \mathbf{x} + (1-\theta)\mathbf{y} \in S \cap \ell \subseteq S$. Because we arbitrarily chose $\mathbf{x}$ and $\mathbf{y}$, we have satisfied the condition of convexity for any two points $\mathbf{x}, \mathbf{y} \in S$, and thus S is a convex set.

Problem 3 *Are the following sets convex?*

1. *A set of the form* $\{x \in \mathbb{R}^n | \alpha \leq w^T x \leq \beta\}$.

2. *The set of points closer to one point than another (by Euclidean distance), i.e.,*

$$\{x | \mathbf{dist}(x, S) \leq \mathbf{dist}(x, Z)\},$$

where $S, Z \in \mathbb{R}^n$.

1. Suppose we have a set $S = \{x \in \mathbb{R}^n | \alpha \leq w^T x \leq \beta\}$ for some $\alpha, \beta \in \mathbb{R}$, $w \in \mathbb{R}^n$ and choose some $x, y \in S$ and $\theta \in [0, 1]$. Then

$$w^T \cdot \theta x + w^T \cdot (1-\theta) y = \theta(w^T x) + (1-\theta)(w^T y) \geq \theta\alpha + (1-\theta)\alpha = \alpha$$

and similarly

$$w^T \cdot \theta x + w^T \cdot (1-\theta) y = \theta(w^T x) + (1-\theta)(w^T y) \leq \theta\beta + (1-\theta)\beta = \beta.$$

Thus, $\theta x + (1-\theta) y \in S$, so S is convex. Notably this means that the set of solutions to a set of linear inequalities is convex.

2. Assume without loss of generality that S is at the origin, and let the coordinates of $T = (x_1^Z, \cdots, x_n^Z)$. Then for any $x = (x_1, \cdots, x_n) \in \mathbb{R}^n$,

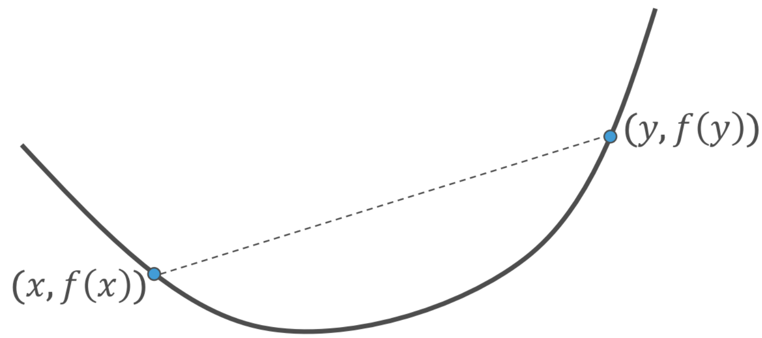
$$\begin{aligned}
\mathbf{dist}(x, S) \leq \mathbf{dist}(x, Z) &\iff \sqrt{\sum_{i=1}^{n} (x_i - 0)^2} \leq \sqrt{\sum_{i=1}^{n} (x_i - x_i^Z)^2} \\
&\iff \sum_{i=1}^{n} x_i^2 \leq \sum_{i=1}^{n} x_i^2 - 2\sum_{i=1}^{n} x_i^Z x_i + \sum_{i=1}^{n} \left(x_1^Z\right)^2 \\
&\iff 0 \leq -2\sum_{i=1}^{n} x_i^Z x_i + \sum_{i=1}^{n} \left(x_i^Z\right)^2 \\
&\iff 2(\boldsymbol{x}^Z)^T \boldsymbol{x} \leq (\boldsymbol{x}^Z)^T (\boldsymbol{x}^Z)
\end{aligned}$$

Since the x_i^Z are constants, this is just a set of linear inequalities, which as shown in part (1) means that the set of points that satisfy these conditions is convex.

1.2 Convex Functions

Definition 3 (Convex Function) *A function* $f : \mathbb{R}^n \to \mathbb{R}$ *is convex if and only if for any* $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ *and* $\theta \in [0, 1]$,

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



For functions that $f : \mathbb{R} \to \mathbb{R}$ that are twice differentiable, this definition is equivalent to saying that $f''(x) \geq 0$ for all $x \in \mathbb{R}$ (this may have been the definition you have seen before in calculus courses!)

Finally, f is concave if and only if $(-1)f$ is convex.

Now let's give some examples for convex functions:

Example 2

- *Exponential.* $f(x) = e^{ax}$. *We can show this is convex via the second derivative:*
$$f''(x) = a^2 e^{ax} \geq 0 \quad \textit{for all } x \in \mathbb{R}.$$
- *Euclidean Norm.* $f(\mathbf{x}) = ||\mathbf{x}||_2 = \sqrt{\sum_{i=1}^{n}(x_i)^2}$. *We can show this is convex via the definition (and via the triangle inequality):*
$$\begin{aligned} ||\theta\mathbf{x} + (1-\theta)\mathbf{y}||_2 &\leq ||\theta\mathbf{x}||_2 + ||(1-\theta)\mathbf{y}||_2 \\ &= \theta||\mathbf{x}||_2 + (1-\theta)||\mathbf{y}||_2 \end{aligned}$$

We also have that weighted sums of convex functions are convex:

Example 3 *Let $f(x) = \sum_{i=1}^{m} a_i f_i(x)$, where f_i is convex and $a_i \geq 0$ for all $i = 1, \ldots, m$. Then f is convex.*

Proof:
$$\begin{aligned} f(\theta x + (1-\theta)y) &= \sum_i a_i f_i(\theta x + (1-\theta)y) \\ &\leq \sum_i a_i(\theta f_i(x) + (1-\theta)f_i(y)) \\ &= \sum_i a_i f_i(x) + (1-\theta)\sum_i a_i f_i(y) \\ &= \theta f(x) + (1-\theta)f(y) \end{aligned}$$

■

Problem 4 *Suppose that $f : \mathbb{R} \to \mathbb{R}$ is increasing and convex on its domain (a, b). Let g denote its inverse, i.e., the function with domain $(f(a), f(b))$ and $g(f(x)) = x$ for $a < x < b$. Suppose that f and g are differentiable. What can you say about the convexity or concavity of g?*

4. Then, choose any $x, y \in (a, b)$ and $\theta \in [0, 1]$. Since f is increasing on its domain, $y > x \iff f(y) > f(x)$. Notably, letting $y' = f(y) \Rightarrow g(y') = y$ and $x' = f(x) \Rightarrow g(x') = x$, this is equivalent to $g(y') > g(x') \iff y' > x'$, so g is also an increasing function. Because of the convexity of f, we have $f(\theta x + (1-\theta)y) \leq \theta f(x) + (1-\theta)f(y)$. Since g is increasing, we can apply it to both sides of the inequality to get

$$\begin{aligned} g(f(\theta x + (1-\theta)y)) &\leq g(\theta f(x) + (1-\theta)f(y)) \\ \theta x + (1-\theta)y &\leq g(\theta x' + (1-\theta)y') \\ \theta g(x') + (1-\theta)g(y') &\leq g(\theta x' + (1-\theta)y') \end{aligned}$$

Since $f : (a, b) \to (f(a), f(b))$ is a bijection, this is true everywhere on the domain of g, and thus g is a concave function.

1.3 Convex Optimization Problems

Let's give a few examples of convex optimization problems.

Example 4 (Linear Programming) *The linear programming problem can be formulated as finding*

$$\min_{\mathbf{x}} \mathbf{c}^T \mathbf{x} \qquad \text{such that } A\mathbf{x} = \mathbf{a} \text{ and } B\mathbf{x} \leq \mathbf{b},$$

where $\mathbf{x} \in \mathbb{R}^n$ is the optimization variable, and $\mathbf{c} \in \mathbb{R}^n$, $A \in \mathbb{R}^{m \times n}$, $\mathbf{a} \in \mathbb{R}^m$, $B \in \mathbb{R}^{k \times n}$, $\mathbf{b} \in \mathbb{R}^k$ are the problem data.

Verbally, we are trying to minimize a (convex) linear objective function subject to linear constraints (which gives us a convex set). Taking the dot product of $\mathbf{x}$ with $\mathbf{c}$ gives us our objective function evaluated at an input of $\mathbf{x}$, the pairing $\mathbf{A}$ and $\mathbf{a}$ encode a set of equality constraints and the pairing $\mathbf{B}$ and $\mathbf{b}$ encode a set of inequality constraints.

Problem 5 *Consider the example of a manufacturer of animal feed who is producing feed mix for dairy cattle. In our simple example the feed mix contains two active ingredients and a filler to provide bulk. One kg of feed mix must contain a minimum quantity of each of four nutrients as below:*

Nutrient	A	B	C	D
gram	90	50	20	2

The ingredients have the following nutrient values and cost:

	A	B	C	D	Cost/kg
Ingredient 1 (gram/kg)	100	80	40	10	40
Ingredient 2 (gram/kg)	200	150	20	0	60

What should be the amounts of active ingredients and filler in one kg of feed mix? Model this as an LP.

5. Let x_1 and x_2 be the number of kilograms of ingredients 1 and 2 respectively that we include in the feed. Then, assuming that we want to minimize cost, the appropriate LP for this situation is

$$\min_{x_1,x_2} 40x_1 + 60x_2$$

such that

$$\begin{aligned}
100x_1 + 200x_2 &\geq 90 \\
80x_1 + 150x_2 &\geq 50 \\
40x_1 + 20x_2 &\geq 20 \\
10x_1 &\geq 2 \\
x_1 + x_2 &\leq 1 \\
x_1, x_2 &\geq 0
\end{aligned}$$

2 Integer Programming

2.1 Introduction

Definition 4 (Feasibility Problem) *Find $x_1, \ldots, x_l$ such that*

$$\begin{aligned}
&\forall i \in [k], \sum_{j=1}^{l} a_{ij} x_j \leq b_i \\
&\forall j \in [l], x_j \in \mathbb{Z}.
\end{aligned}$$

Definition 5 (Integer Programming Optimization Problem)

$$\max \sum_{j=1}^{l} c_j x_j$$

such that

$$\forall i \in [k], \sum_{j=1}^{l} a_{ij} x_j \leq b_i$$
$$\forall j \in [l], x_j \in \mathbb{Z}.$$

Example 5 (Envy-Free) *Suppoes we have players $N = \{1, \ldots, n\}$ and items $M = \{1, \ldots, M\}$, and player i has value v_{ij} for item j. Partition the items into bundles $A_1, \ldots, A_n$. We say that the $A_1, \ldots, A_n$ is envy-free iff*

$$\forall i, i', \sum_{j \in A_i} v_{ij} \geq \sum_{j \in A_{i'}} v_{ij}.s$$

Formulate this as an IP.

Proof: Let $x_{ij} \in \{0, 1\}$, $x_{ij} = 1$ iff $j \in A_i$. Then our feasibility problem is:

$$\begin{aligned}
&\text{Find } x_{11}, \ldots, x_{nm} \text{ such that} \\
&\forall i \in N, \forall i' \in N, \sum_{j \in M} v_{ij} x_{ij} \geq \sum_{j \in M} v_{ij} x_{i'j} \\
&\forall j \in M, \sum_{i \in N} x_{ij} = 1 \\
&\forall i \in N, j \in M, x_{ij} \in \{0, 1\}
\end{aligned}$$

■

Problem 6 *Recall the 8 queens puzzle: on an 8×8 grid, place 8 queens such that no two are in the same row, column, or diagonal. Formulate this as an integer program.*

6. Let $x_{i,j}$ be the indicator variable for whether there is a queen in the ith row and jth column, i.e. it equals 1 if there is a queen in cell (i, j) and 0 otherwise. Then, we can formulate each of the conditions as enforcing that the sum over each of the rows, columns, and diagonals equal 1, ensuring that there is exactly one queen in that set of grid cells.

Find

$$x_{i,j} \;\; \forall i, j \in \{1, 2, \cdots, 8\}$$

such that

$$
\begin{aligned}
x_{ij} &\in \{0,1\} && \forall i,j \in \{1,2,\cdots,8\} \\
\sum_{i=1}^{8} x_{i,j} &\leq 1 && \forall j \in \{1,2,\cdots,8\} && \text{(columns)} \\
\sum_{j=1}^{8} x_{i,j} &\leq 1 && \forall i \in \{1,2,\cdots,8\} && \text{(rows)} \\
\sum_{i=\max(1,k-8)}^{\min(8,k-1)} x_{i,k-i} &\leq 1 && \forall k \in \{2,3,\cdots,16\} && \text{(diagonals from bottom left to top right)} \\
\sum_{i=\max(1,1+k)}^{\min(8,k+8)} x_{i,i+k} &\leq 1 && \forall k \in \{-7,-6,\cdots,6,7\} && \text{(diagonals from top left to bottom right)} \\
\sum_{i=1}^{8}\sum_{j=1}^{8} x_{ij} &= 8 && && \text{(8 total queens)}
\end{aligned}
$$