

# Convex Analysis

Dirk Lorenz

## Contents

<b>1</b>	<b>Convex sets</b>	<b>3</b>
<b>2</b>	<b>Hyperplanes and cones</b>	<b>6</b>
<b>3</b>	<b>Projection and separation</b>	<b>11</b>
<b>4</b>	<b>Convex functions</b>	<b>16</b>
<b>5</b>	<b>Characterization of convex functions</b>	<b>20</b>
<b>6</b>	<b>Continuity of convex functions and minimizers</b>	<b>24</b>
<b>7</b>	<b>Inf-projection and inf-convolution</b>	<b>27</b>
<b>8</b>	<b>Proximal mappings and Moreau-Yoshida regularization</b>	<b>31</b>
<b>9</b>	<b>Subgradients</b>	<b>35</b>
<b>10</b>	<b>Subdifferential calculus</b>	<b>40</b>
<b>11</b>	<b>Applications of subgradients</b>	<b>44</b>
<b>12</b>	<b>Proximal algorithms</b>	<b>50</b>
<b>13</b>	<b>Convex conjugation</b>	<b>55</b>
<b>14</b>	<b>Conjugation calculus</b>	<b>60</b>
<b>15</b>	<b>Fenchel-Rockafellar duality</b>	<b>65</b>
<b>16</b>	<b>Examples of duality and optimality systems</b>	<b>70</b>
<b>17</b>	<b>Classes of optimization problems</b>	<b>75</b>
<b>18</b>	<b>Convergence rates and worst case analysis</b>	<b>80</b>
<b>19</b>	<b>More worst case analysis</b>	<b>84</b>
<b>20</b>	<b>Subgradient method and gradient descent</b>	<b>89</b>
<b>21</b>	<b>Accelerated gradient descent</b>	<b>93</b>

<b>22 Analysis of the proximal gradient method and it's acceleration</b>	<b>99</b>
<b>23 Monotone operators</b>	<b>103</b>
<b>24 Resolvents and non-expansive operators</b>	<b>108</b>
<b>25 Relaxed Mann iterations</b>	<b>111</b>
<b>26 The proximal point algorithm</b>	<b>115</b>
<b>27 Preconditioned proximal point methods</b>	<b>118</b>
<b>28 Primal dual methods</b>	<b>119</b>

## 1 Convex sets

We introduce the basic notion of convexity of sets. We will develop everything in the euclidean space  $\mathbb{R}^d$ , but most of what we will be doing, will also work in a general real and separable Hilbert space  $X$ , i.e. a real vector space that is equipped with an inner product and has a orthonormal basis.

**Definition 1.1.** A set  $C \subset \mathbb{R}^d$  is *convex*, if for all  $x, y \in C$  and  $\lambda \in [0, 1]$  it holds that  $\lambda x + (1 - \lambda)y \in C$ . The term  $\lambda x + (1 - \lambda)y$  is called the *convex combination*.

More general: For  $x_1, \dots, x_n \in \mathbb{R}^d$  and  $\lambda_1, \dots, \lambda_n \geq 0$  with  $\sum_{i=1}^n \lambda_i = 1$  we call  $x = \sum_{i=1}^n \lambda_i x_i$  a *convex combination*.

**Definition 1.2.** The *convex hull*  $\text{conv}(S)$  of a subset  $S \subset \mathbb{R}^d$  is the set of all convex combinations of points in  $S$ .

**Example 1.3.** For two convex sets  $C_1, C_2 \subset \mathbb{R}^d$  it holds that the *Minkowski sum*

$$C_1 + C_2 := \{x = x_1 + x_2 \mid x_1 \in C_1, x_2 \in C_2\}$$

is again convex.  $\triangle$

It's fairly straightforward to prove the following (and one should do so as an exercise):

**Proposition 1.4.** The following sets are convex:

1.  $\alpha C = \{\alpha x \mid x \in C\}$ , for convex  $C \in \mathbb{R}^d$  and  $\alpha \in \mathbb{R}$ ,
2.  $AC \subset \mathbb{R}^m$ , for convex  $C \in \mathbb{R}^d$  and a matrix  $A \in \mathbb{R}^{m \times d}$ ,
3.  $C_1 \times C_2 \subset \mathbb{R}^{m+d}$  for convex  $C_1 \in \mathbb{R}^m$  and convex  $C_2 \in \mathbb{R}^d$ ,
4.  $\bigcap_{i \in I} C_i$  for convex  $C_i$  and any index set  $I$ ,
5. the closure  $\bar{C}$  and the interior  $C^\circ$  for convex  $C$ .

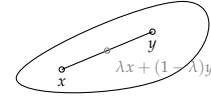
*Proof.* Let us prove the second point: If  $x, y$  are in  $AC$  then there are  $u, v$ , such that  $x = Au$  and  $y = Av$ . For any  $\lambda \in [0, 1]$  it holds that  $\lambda u + (1 - \lambda)v \in C$ , and hence  $\lambda x + (1 - \lambda)y = \lambda Au + (1 - \lambda)Av = A(\lambda u + (1 - \lambda)v)$  is in  $AC$ .  $\square$

The proofs of the remaining assertions are straightforward and left as exercise.

**Definition 1.5.** A set  $S \subset \mathbb{R}^d$  is called *affine* if for all  $x, y \in S$  and  $\lambda \in \mathbb{R}$  it holds that  $\lambda x + (1 - \lambda)y \in S$ .

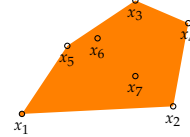
Of course, linear subspaces are also affine spaces and for every non-empty affine set, there is a unique subspace  $L$  and some vector  $a$  such that  $S = a + L$ . It's also clear that affine sets are convex.

A point  $x$  is an *affine combination* of  $x_1, \dots, x_n$  if there exists  $\lambda_i$  with  $\sum_{i=1}^n \lambda_i = 1$  and  $x = \sum_{i=1}^n \lambda_i x_i$ .



It holds: A set is convex if and only if it contains all convex combinations of its points.

Another way to say this is: The convex hull is the smallest convex set that contains  $S$ .



$S = \{x_1, \dots, x_7\}$

We will denote the closure of  $C$  also by  $\text{cl}(C)$  and the interior also by  $\text{int}(C)$ .

Note that  $\lambda x + (1 - \lambda)y = y + \lambda(x - y)$ , i.e. the point  $\lambda x + (1 - \lambda)y$  is reached by starting from  $y$  and going  $\lambda$  times the vector from  $y$  to  $x$  in the direction of  $x$ .

**Definition 1.6.** The *affine hull*  $\text{aff}(S)$  of a set  $S$  is the set of all affine combinations of elements of  $S$ .

Alternatively, we could also define the affine hull of  $S$  as the smallest affine set which contains  $S$ .

An affine space inherits a topology from the surrounding space  $\mathbb{R}^d$  and hence, we have closure and interior with respect to this topology for subsets of affine spaces. This gives rise to the following notion:

**Definition 1.7.** The *relative interior* of some  $S \subset \mathbb{R}^d$ , denoted by  $\text{ri}(S)$ , is the interior of  $S$  relative to the affine set  $\text{aff}(S)$ , i.e.

$$\text{ri}(C) = \{x \in C \mid \exists \epsilon > 0 : B_\epsilon(x) \cap \text{aff}(C) \subset C\}.$$

The set  $\overline{C} \setminus \text{ri}(C)$  is called *relative boundary* of  $C$ .

**Proposition 1.8.** 1. For convex  $C \neq \emptyset$  it holds that  $\text{ri}(C)$  is also convex and it holds that  $\text{aff}(\text{ri}(C)) = \text{aff}(\overline{C})$ .

2. For convex  $C$ , invertible  $A \in \mathbb{R}^{d \times d}$  and any  $b \in \mathbb{R}^d$  it holds that  $A \text{ri}(C) + b = \text{ri}(AC + b)$  and for all  $A \in \mathbb{R}^{m \times d}$  it holds that  $A(\text{ri}(C)) = \text{ri}(AC)$ .
3. For convex  $C$  we have  $x \in \text{ri}(C)$  if and only if for every  $y \in \text{aff}(C)$  there exists  $\epsilon > 0$  such that  $x \pm \epsilon(y - x) \in C$ ,
4. For convex  $C_1, C_2$  it holds  $\overline{C_1} = \overline{C_2}$  if and only if  $\text{ri}(C_1) = \text{ri}(C_2)$ ,
5. For convex  $C_1, C_2$  it holds  $\text{ri}(C_1 + C_2) = \text{ri}(C_1) + \text{ri}(C_2)$ .

You also see  $\text{relint}(S)$  for the relative interior. Note that there is no notion for relative closure as the closure is always within the affine hull.

We don't give a proof of this proposition, but note that point 2. is helpful to prove other statements about the relative interior: If  $\text{aff}(C)$  is an  $m$ -dimensional affine space, we can, without loss of generality, assume that  $\text{aff}(C)$  lies in the subspace  $V = \{x \mid x_{m+1} = \dots = x_n = 0\}$  and since this is just a copy of  $\mathbb{R}^m$ , we can assume that  $C$  is full-dimensional, i.e.  $\text{aff}(C)$  is the full space.

Note that even if  $C_1 \subset C_2$ , it is generally not true, that  $\text{ri}(C_1)$  is contained in  $\text{ri}(C_2)$ ! This may be seen, for example, with  $C_2$  being a (closed) square in  $\mathbb{R}^2$  and  $C_1$  being one of its sides.

**Definition 1.9.** A set of  $n + 1$  points  $x_0, \dots, x_n \in \mathbb{R}^d$  are called *affinely independent* if the affine hull  $\text{aff}(\{x_0, \dots, x_n\})$  is an  $n$ -dimensional affine space.

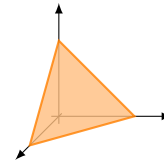
Since  $\text{aff}\{x_0, \dots, x_n\} = V + x_0$  where  $V$  is the subspace spanned by the vectors  $x_1 - x_0, \dots, x_n - x_0$ , we see that the vectors  $x_0, \dots, x_n$  are affinely independent if and only if the vectors  $x_1 - x_0, \dots, x_n - x_0$  are linearly independent.

**Proposition 1.10.** If  $x_0, \dots, x_n$  are affinely independent, each  $x \in \text{aff}\{x_0, \dots, x_n\}$  can be represented uniquely as an affine combination  $x = \sum_{i=0}^n \lambda_i x_i$  and these  $\lambda_i$  are called *barycentric coordinates* of  $x$  with respect to the points  $x_0, \dots, x_n$ .

*Proof.* If  $M = \text{aff}\{x_0, \dots, x_n\} = x_0 + V$  with  $V = \text{span}\{x_1 - x_0, \dots, x_n - x_0\}$ , we can express each  $y \in V$  uniquely as  $y = \sum_{i=1}^n \lambda_i (x_i - x_0)$  and since each  $x \in M$  is of the form  $x_0 + y$  with  $y \in V$ , we express each  $x \in M$  uniquely as  $x = \sum_{i=1}^n \lambda_i (x_i - x_0) + x_0$ , i.e. as an affine combination  $x = \sum_{i=0}^n \lambda_i x_i$  with  $\sum_{i=0}^n \lambda_i = 1$ .  $\square$

If we have affinely independent points  $x_0, \dots, x_n$  their convex hull is called a *simplex*. Of special importance is the *probability simplex* (also called *standard simplex*) which is the convex hull of the standard basis vectors  $e_i$ , i.e.

$$\Delta_d := \text{conv}(e_1, \dots, e_d)$$



## 2 Hyperplanes and cones

For each  $x_0, p \in \mathbb{R}^d$  with  $p \neq 0$  the *hyperplane* through  $x_0$  with normal vector  $p$  can be written with  $\alpha = \langle x_0, p \rangle$  as

$$H_{p,\alpha} := \{x \in \mathbb{R}^d \mid \langle p, x \rangle = \alpha\} = \{x \mid \langle p, x - x_0 \rangle = 0\}.$$

Hyperplanes are affine sets and the orthogonal projection of some  $x$  onto  $H_{p,\alpha}$  is

$$\hat{x} = x - \frac{\langle p, x \rangle - \alpha}{\|p\|_2^2} p.$$

Moreover, there are the associated half-spaces

$$H_{p,\alpha}^+ := \{x \in \mathbb{R}^d \mid \langle p, x \rangle \geq \alpha\}, \quad H_{p,\alpha}^- := \{x \in \mathbb{R}^d \mid \langle p, x \rangle \leq \alpha\}.$$

We will show (with the help of separation theorems) that a closed and convex set equals the intersection of all half-spaces that contain  $C$ .

A set  $C$  that is the intersection of finitely many half-spaces is called *polyhedral set*, in this case we can write

$$C = \{x \mid Ax \leq b, Bx = d\}$$

for some  $A \in \mathbb{R}^{n \times d}, b \in \mathbb{R}^n, B \in \mathbb{R}^{m \times d}, c \in \mathbb{R}^m$ .

Related to hyperplanes are *affine functionals* which are of the form

$$f(x) = \langle p, x \rangle + \alpha$$

for some  $p \in \mathbb{R}^d, \alpha \in \mathbb{R}$ . A hyperplane in  $\mathbb{R}^{d+1}$  is of the form

$$H_{p,\alpha} = \{(x', x_{d+1}) \in \mathbb{R}^{d+1} \mid \langle p', x' \rangle + p_{d+1} x_{d+1} = \alpha\}.$$

A “vertical” hyperplane is one with  $p_{d+1} = 0$  but for the other (i.e. for  $p_{d+1} \neq 0$  we have

$$x_{d+1} = \left\langle -\frac{p'}{p_{d+1}}, x' \right\rangle + \frac{\alpha}{p_{d+1}},$$

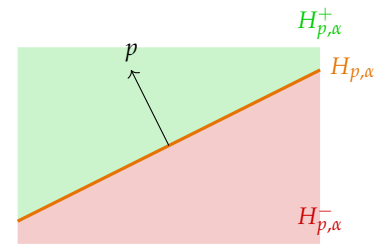
i.e., the hyperplane  $H_{p,\alpha} \subset \mathbb{R}^{d+1}$  is the graph of the affine function  $f: \mathbb{R}^d \rightarrow \mathbb{R}, f(x') = \left\langle -\frac{p'}{p_{d+1}}, x' \right\rangle + \frac{\alpha}{p_{d+1}}$ .

**Definition 2.1.** A set  $K \subset \mathbb{R}^d$  is called a *cone*, if  $x \in K$  and  $\lambda \geq 0$  implies  $\lambda x \in K$ .

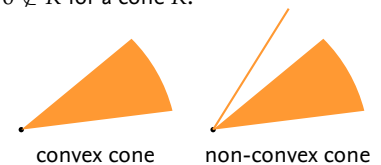
**Proposition 2.2.** A cone  $K$  is convex if and only if  $K + K \subset K$ .

*Proof.* “ $\Rightarrow$ ”: If  $K$  is convex and  $x, y \in K$  we have  $(x + y)/2 \in K$  and hence  $x + y \in K$ .

“ $\Leftarrow$ ”: Let  $K + K \subset K$  hold and let  $x, y \in K$ , since  $K$  is a cone, we have  $\lambda x, (1 - \lambda)x \in K$  for any  $\lambda \geq 0$  and hence,  $\lambda x + (1 - \lambda)y \in K$  by  $K + K \subset K$ .  $\square$



Sometimes cones are defined with just  $\lambda > 0$ . In this case, one may have  $0 \notin K$  for a cone  $K$ .



- Example 2.3.**
1. Every linear subspace is a convex cone, more precisely, a cone  $K$  is a linear subspace if and only if  $K = -K$ .
  2. The halfspaces  $H_{p,0}^+ = \{x \mid \langle p, x \rangle \geq 0\}$  are convex cones.
  3. An important convex cone is the non-negative orthant  $K = \mathbb{R}_{\geq 0}^d \{x \mid x_1, \dots, x_n \geq 0\}$ .
  4. The set  $\{Ay \mid y \geq 0\}$  is a convex cone for  $A \in \mathbb{R}^{d \times m}$  and cones of this form are called *finitely generated*.
  5. The set  $\{x \mid A^T x \leq 0\}$  with  $A \in \mathbb{R}^{d \times m}$  is a convex cone and cones of this type are called *polyhedral cones*.
  6. The set

$$\mathbb{L}^{d+1} = \{(x', x_{d+1}) \in \mathbb{R}^{d+1} \mid \|x'\|_2 \leq x_{d+1}\}$$

is a convex cone called *second order cone* (or Lorentz-cone or, due to its shape, ice-cream cone).

7. The set of symmetric positive semi-definite matrices  $\text{Sym}_d^+ = \{M \in \mathbb{R}^{d \times d} \mid M \succcurlyeq 0\}$  is a convex cone. For  $d = 2$  symmetric matrices are of the form

$$M = \begin{bmatrix} x_1 & x_2 \\ x_2 & x_3 \end{bmatrix}$$

and such a matrix is positive semi-definite if  $x_1 \geq 0$  and  $x_1 x_3 \geq x_2^2$ . Using the change of variables  $x = x_1 - x_3$ ,  $y = 2x_2$ ,  $z = x_1 + x_3$  one can see that  $M$  is positive semi-definite if and only if  $(x, y, z) \in \mathbb{L}^3$ .

△

**Definition 2.4.** For a non-empty set  $S$  one defines the *polar cone* by

$$S^* := \{p \in \mathbb{R}^d \mid \forall x \in S : \langle p, x \rangle \leq 0\} = \bigcap_{x \in S} H_{x,0}^-.$$

Consequently,  $S^*$  is always closed and convex. Since for all  $x \in S$  and  $p \in S^*$  we always have

$$\frac{\langle p, x \rangle}{\|p\|_2 \|x\|_2} \leq 0,$$

and the left hand side defines the cosine of the angle between  $x$  and  $p$ , we see that this angle is always larger or equal to  $\pi/2$ .

One can see that  $(K^*)^* = \overline{\text{conv } K}$  for any set  $K$  and hence, for closed convex cones  $K$  it holds that  $(K^*)^* = K$ .

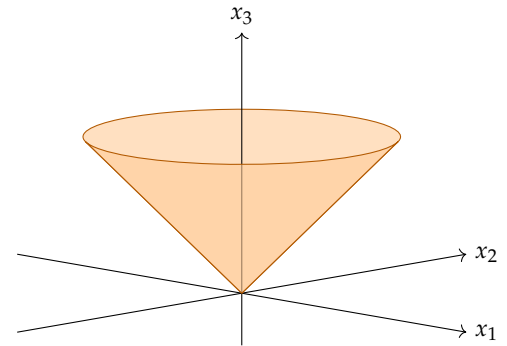
**Example 2.5.**

1. The polar cone of  $K = \{0\}$  is  $K^* = \mathbb{R}^d$ .

2. If  $K$  is a linear subspace, one has  $K^* = K^\perp$ .

We also write  $x \geq 0$  if every component is non-negative.

The Lorentz cone  $\mathbb{L}^3$ :



3. For the non-negative orthant  $\mathbb{R}_{\geq 0}^d$  the polar cone is  $(\mathbb{R}_{\geq 0}^d)^* = \mathbb{R}_{\leq 0}^d$ , i.e. the non-positive orthant.
4. For some  $p \in \mathbb{R}^d$  and  $K = H_{p,0}^- = \{x \mid \langle p, x \rangle \leq 0\}$  it holds that  $K^* = \{\lambda p \mid \lambda \geq 0\}$ .
5. The polar cone of the finitely generated cone  $K = \{Ay \mid y \geq 0\}$  for some  $A \in \mathbb{R}^{d \times m}$  is the polyhedral cone  $K^* = \{p \mid A^T p \leq 0\}$ .

△

**Definition 2.6.** For some set  $S \subset \mathbb{R}^d$  and  $x_0 \in S$  we define the *tangential cone* of  $S$  in  $x_0$  by

$$T_S(x_0) := \{v \in \mathbb{R}^d \mid \exists x_n \in S, x_n \rightarrow x_0, \lambda_n \searrow 0 : v = \lim_{n \rightarrow \infty} (x_n - x_0) / \lambda_n\}$$

and the *normalized directions* in  $T_S(x_0)$  are the limits

$$\frac{x_n - x_0}{\|x_n - x_0\|_2} \rightarrow \frac{v}{\|v\|_2}.$$

If  $C$  is convex, one has

$$T_C(x_0) = \overline{\{t(x - x_0) \mid x \in C, t \geq 0\}}$$

which shows that  $T_C(x_0)$  is closed.

**Definition 2.7.** For a convex  $C$  and  $x_0 \in C$  we define the *normal cone* of  $C$  at  $x_0$  as

$$N_C(x_0) := \{p \mid \forall x \in C : \langle p, x - x_0 \rangle \leq 0\}.$$

By definition, the normal cone is always closed.

**Theorem 2.8.** For a convex set  $C$  and  $x_0 \in C$  it holds that

$$(T_C(x_0))^* = N_C(x_0).$$

*Proof.* Let  $p \in T_C(x_0)^*$ . Then it holds that  $\langle p, v \rangle \leq 0$  for all  $v \in T_C(x_0)$ , i.e.  $\langle p, x - x_0 \rangle \leq 0$ .

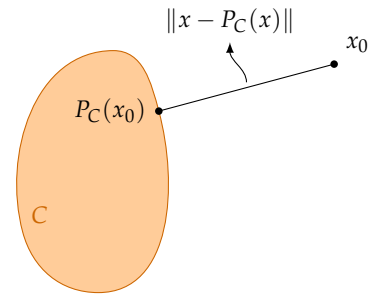
Conversely, let  $p \in N_C(x_0)$ . We have to show that  $\langle p, v \rangle \leq 0$  for any  $v \in T_C(x_0)$ . Since we can write  $v = \lim_{n \rightarrow \infty} (x_n - x_0) / \lambda_n$  with  $x_n \rightarrow x_0$ ,  $x_n \in C$  and  $\lambda_n \searrow 0$ , we have  $\langle p, x_n - x_0 \rangle \leq 0$  and hence  $\langle p, v \rangle \leq 0$  as well. □

The following theorem would better fit in the next section, however, we still have some space here, so we state and prove it now:

**Theorem 2.9** (Projection theorem). Let  $C \subset \mathbb{R}^d$  be a nonempty, closed, convex set. Then, for any  $x_0 \in \mathbb{R}^d$ , there exists a unique  $\hat{x}_0 \in C$ , called *orthogonal projection* of  $x_0$  onto  $C$ , and denoted by  $P_C x_0 := \hat{x}_0$ , such that

$$\|x_0 - \hat{x}_0\|_2 = \inf_{x \in C} \|x_0 - x\|_2$$





and this element fulfills

$$\forall x \in C : \langle x_0 - \hat{x}_0, x - \hat{x}_0 \rangle \leq 0. \quad (1)$$

Conversely, if  $y \in C$  fulfills the variational inequality

$$\forall x \in C : \langle x_0 - y, x - y \rangle \leq 0, \quad (2)$$

then  $y = P_C x_0$ .

*Proof.* The function  $f(x) = \|x - x_0\|_2$  is continuous and since  $C$  is closed,  $f$  takes its infimum in  $C$  (in fact, in the compact set  $C \cap B_r(x_0)$  for some large  $r$ ). By Weierstraß' theorem, the infimum is attained. This shows existence of a projection.

Now let  $\hat{x}_0$  be an orthogonal projection of  $x_0$  onto  $C$ . By convexity of  $C$ , we have  $\hat{x}_0 + \lambda(x - \hat{x}_0) \in C$  for every  $x \in C$  and  $\lambda \in [0, 1]$ . Since  $f$  is minimal at  $\hat{x}_0$ , the same holds for  $x \mapsto \|x - x_0\|_2^2$  which means

$$\begin{aligned} 0 &\leq \|\hat{x}_0 + \lambda(x - \hat{x}_0)\|_2^2 - \|\hat{x}_0 - x_0\|_2^2 \\ &= \lambda^2 \|x - \hat{x}_0\|_2^2 - 2\lambda \langle \hat{x}_0 - x_0, \hat{x}_0 - x \rangle. \end{aligned}$$

For  $\lambda > 0$  we can rearrange to

$$\langle \hat{x}_0 - x_0, \hat{x}_0 - x \rangle \leq \frac{\lambda}{2} \|x - \hat{x}_0\|_2^2$$

and with  $\lambda \rightarrow 0$  we have shown that (1) holds.

Conversely, assume that the variational inequality (2) holds for some  $y$ . Then, by Cauchy-Schwarz, it holds for all  $x \in C$  that

$$\begin{aligned} 0 &\geq \langle y - x_0, y - x \rangle = \langle y - x_0, y - x_0 + x_0 - x \rangle \\ &= \|y - x_0\|_2^2 + \langle y - x_0, x_0 - x \rangle \\ &\geq \|y - x_0\|_2^2 - \|y - x_0\|_2 \|x - x_0\|_2. \end{aligned}$$

Dividing by  $\|y - x_0\|_2$  we obtain  $\|y - x_0\|_2 \leq \|x - x_0\|_2$  for all  $x \in C$  and this says that  $y$  is the orthogonal projection of  $x_0$  onto  $C$ .

Finally, we show uniqueness: Assume that  $\hat{x}_0$  and  $\tilde{x}_0$  are both orthogonal projections of  $x_0$  onto  $C$ . Since  $\hat{x}_0, \tilde{x}_0 \in C$  we can plug them into their variational inequalities and get

$$\langle x_0 - \hat{x}_0, \tilde{x}_0 - \hat{x}_0 \rangle \leq 0, \quad \langle x_0 - \tilde{x}_0, \hat{x}_0 - \tilde{x}_0 \rangle \leq 0.$$

Adding both inequalities we get  $\|\hat{x}_0 - \tilde{x}_0\|_2^2 \leq 0$  which means  $\hat{x}_0 = \tilde{x}_0$ .  $\square$

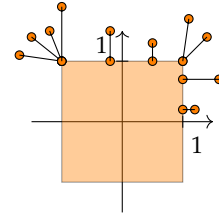
*Example 2.10.* We project onto balls in the  $p$ -norms for  $p = 1, 2, \infty$ :

1. First consider  $p = \infty$  and the respective norm ball of radius  $\lambda > 0$  around 0 is

$$B_\lambda^\infty(0) := \{x \mid \max(|x_1|, \dots, |x_d|) \leq \lambda\}.$$

The projection of some  $x$  onto this ball is minimizing  $f(x - y) = \|x - y\|_2$  over all  $y \in B_\lambda^\infty(0)$ , i.e. over all  $y$  with  $|y_i| \leq \lambda$ . This can be done componentwise and leads to

$$(P_{B_\lambda^\infty(0)}x)_i = \begin{cases} x_i, & \text{if } |x_i| \leq \lambda \\ \lambda \operatorname{sign}(x_i), & \text{if } |x_i| > \lambda \end{cases}$$



which can be written consisely by  $P_{B_\lambda^\infty(0)}x = \min(\max(x, -\lambda), \lambda)$  where the minimum and maximum applied componentwise.

- For  $p = 2$  we simply need to shrink  $x$  if it is outside of the ball, i.e.

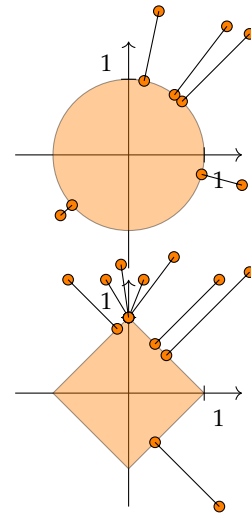
$$P_{B_\lambda^2(0)}x = \begin{cases} x, & \text{if } \|x\|_2 \leq \lambda \\ \lambda \frac{x}{\|x\|_2}, & \text{if } \|x\|_2 > \lambda. \end{cases} = \max(1, \frac{\lambda}{\|x\|_2})x$$

- The case  $p = 1$  is more complicated and there is no explicit formula. However, one can show the following: We define the *soft-shrinkage* (or soft-thresholding) function

$$S_\lambda(x) = \max(|x| - \lambda, 0) \operatorname{sign}(x)$$

and denote by  $\pi$  a permutation of  $\{1, \dots, d\}$  which sorts the entries of  $x$  in decreasing order, i.e.  $|x_{\pi(1)}| \geq |x_{\pi(2)}| \geq \dots \geq |x_{\pi(d)}| \geq 0$ . Then, if  $m$  is the largest index such that  $|x_{\pi(m)}| > 0$  and  $\frac{|x_{\pi(1)}| + \dots + |x_{\pi(m)}| - \lambda}{m} \leq |x_{\pi(m)}|$ , one has

$$P_{B_\lambda^1(0)}x = \begin{cases} x, & \text{if } \|x\|_1 \leq \lambda \\ S_\mu(x), & \text{if } \|x\|_1 > \lambda. \end{cases}$$



△

### 3 Projection and separation

**Theorem 3.1.** Let  $K \subset \mathbb{R}^d$  be a non-empty, closed and convex cone. Then every element in  $x_0$  can be uniquely decomposed as

$$x_0 = P_K x_0 + P_{K^*} x_0$$

and it holds that  $P_K x_0 \perp P_{K^*} x_0$ .

*Proof.* By the Projection Theorem (Theorem 2.9) one has for every  $x \in K$

$$\langle x_0 - P_K x_0, x - P_K x_0 \rangle \leq 0. \quad (*)$$

For  $x = 0$  we have  $\langle x_0 - P_K x_0, P_K x_0 \rangle \geq 0$  and for  $x = 2P_K x_0 \in K$  one has  $\langle x_0 - P_K x_0, P_K x_0 \rangle \leq 0$  which implies that we have

$$\langle x_0 - P_K x_0, P_K x_0 \rangle = 0. \quad (**)$$

Thus, by (\*) we have for all  $x \in K$

$$\langle x_0 - P_K x_0, x \rangle \leq 0$$

and this means that  $x_0 - P_K x_0 \in K^*$  by the definition of the polar cone. Moreover, since for all  $x \in K^*$

$$\begin{aligned} \langle x_0 - (x_0 - P_K x_0), x - (x_0 - P_K x_0) \rangle &= \langle P_K x_0, x - x_0 + P_K x_0 \rangle \\ &= \langle P_K x_0, x \rangle + \langle P_K x_0, P_K x_0 - x_0 \rangle \leq 0 \end{aligned}$$

we have (again by the Projection Theorem) that  $x_0 - P_K x_0 = P_{K^*} x_0$ . The orthogonality follows from (\*\*).  $\square$

Now we come to the notion of separation:

**Definition 3.2.** Let  $C_1, C_2$  be two sets. We say that a hyperplane  $H_{p,\alpha}$

1. *separates*  $C_1$  and  $C_2$  if for all  $x_1 \in C_1$  and  $x_2 \in C_2$  it holds that

$$\langle p, x_1 \rangle \leq \alpha \leq \langle p, x_2 \rangle.$$

2. *strictly separates*  $C_1$  and  $C_2$  if for all  $x_1 \in C_1$  and  $x_2 \in C_2$  it holds that

$$\langle p, x_1 \rangle < \alpha < \langle p, x_2 \rangle.$$

3. *properly separates*  $C_1$  and  $C_2$  if it separates the sets and there exist  $x_i \in C_i, i = 1, 2$  such that

$$\langle p, x_1 \rangle < \langle p, x_2 \rangle.$$

Since the polar cone of a subspace  $V$  is the orthogonal complement  $V^\perp$  we obtain:

**Corollary.** Let  $V$  be a non-empty subspace of  $\mathbb{R}^d$ . Then the orthogonal projection of  $x_0$  onto  $V$  is characterized by

$$\forall x \in V : \langle x_0 - P_V x_0, x \rangle = 0$$

and, moreover,  $x_0 = P_V x_0 + P_{V^\perp} x_0$ .

In terms of halfspaces:  $C_1 \subset H_{p,\alpha}^-$  and  $C_2 \in H_{p,\alpha}^+$ .

In terms of halfspaces:  $C_1 \subset \text{int}(H_{p,\alpha}^-)$  and  $C_2 \subset \text{int}(H_{p,\alpha}^+)$ .

**Theorem 3.3.** If  $C$  is a non-empty, closed and convex set and  $x_0 \notin C$ , then we can strictly separate  $C$  from  $\{x_0\}$ , i.e. there exists  $p$  and  $\epsilon > 0$  such that

$$\sup_{x \in C} \langle p, x \rangle \leq \langle p, x_0 \rangle - \epsilon.$$

*Proof.* By the Projection Theorem (Theorem 2.9) we have for all  $x \in C$  that

$$\langle P_C x_0 - x_0, P_C x_0 - x \rangle \leq 0,$$

from which we deduce by adding and subtracting  $x_0$  in the right argument that

$$\|P_C x_0 - x_0\|_2^2 \leq \langle x_0 - P_C x_0, x_0 - x \rangle$$

for all  $x \in C$ . But since  $x_0 \notin C$ , we have  $\|P_C x_0 - x_0\|_2^2 \geq \epsilon > 0$  so we can take  $p = x_0 - P_C x_0$ .  $\square$

**Corollary 3.4.** For a closed and convex set  $C$  it holds that  $C$  equals the intersection of all halfspaces that include  $C$ .

The case of  $C = \emptyset$  is clear. That  $C$  is a subset of said intersection is clear. And if  $x \notin C$  we can find a hyperplane that separates  $C$  from  $x$  and hence, there is a corresponding halfspace that contains  $C$ , but not  $x$  and hence,  $x$  is not in said intersection.

**Theorem 3.5** (Strong separation). Let  $C_1, C_2$  be non-empty, convex and disjoint and let  $C_1$  be closed and  $C_2$  be compact. Then there exists  $p \in \mathbb{R}^d$  and  $\alpha \in \mathbb{R}$  such that  $H_{p,\alpha}$  strictly separates  $C_1$  and  $C_2$ , i.e. for all  $x_1 \in C_1, x_2 \in C_2$  it holds that

$$\langle p, x_1 \rangle < \alpha < \langle p, x_2 \rangle.$$

*Proof.* We consider the Minkowski sum  $C := C_1 + (-C_2) = \{x_1 - x_2 \mid x_1 \in C_1, x_2 \in C_2\}$  which is also non-empty, closed and convex. Also  $0 \notin C$  since  $C_1$  and  $C_2$  are disjoint. Hence, we can consider  $\hat{x} = P_C 0 \in C$ . Since this is a point in  $C = C_1 - C_2$  we can write it as  $\hat{x} = \hat{x}_1 - \hat{x}_2$  with  $x_i \in C_i, i = 1, 2$ . Now we set

$$x^* := \frac{\hat{x}_2 + \hat{x}_1}{2}, \quad p := \frac{\hat{x}_2 - \hat{x}_1}{2}, \quad \alpha := \langle p, x^* \rangle.$$

Note that  $p \neq 0$ . By the Projection Theorem, we get for all  $x_1 \in C_1, x_2 \in C_2$  that

$$\langle 0 - (\hat{x}_1 - \hat{x}_2), x_1 - x_2 - (\hat{x}_1 - \hat{x}_2) \rangle \leq 0$$

from which we deduce (using  $x^* - \hat{x}_1 = (\hat{x}_2 - \hat{x}_1)/2$  and  $x^* - \hat{x}_2 = (\hat{x}_1 - \hat{x}_2)/2$ )

$$\langle x^* - \hat{x}_1, x_1 - \hat{x}_1 \rangle + \langle x^* - \hat{x}_2, x_2 - \hat{x}_2 \rangle \leq 0.$$

Plugging in  $x_1 = \hat{x}_1$  and  $x_2 = \hat{x}_2$ , respectively, we get for all  $x_i \in C_i$  that

$$\langle x^* - \hat{x}_i, x_i - \hat{x}_i \rangle \leq 0.$$

This means that  $\hat{x}_i$  is the orthogonal projection of  $x^*$  onto  $C_i$ . Finally, since  $x^* - \hat{x}_1 = p$  we get

$$\langle p, x_1 \rangle \leq \langle p, \hat{x}_1 \rangle = \langle p, x^* \rangle + \langle p, \hat{x}_1 - x^* \rangle = \alpha - \|p\|_2^2 < \alpha.$$

Similarly, (using  $x^* - \hat{x}_2 = -p$ ) one shows that  $\langle p, x_2 \rangle > \alpha$ .  $\square$

We recall the following fact from analysis:

**Proposition 3.6.** *If  $(S_x)_x$  is a family of compact subsets of a metric space such that the intersection of every finity subfamily of the  $S_x$  is non-empty, then  $\bigcap_x S_x \neq \emptyset$ .*

**Proposition 3.7.** *Let  $C$  be a non-empty convex set and  $x_0 \notin C^\circ$ . Then there exists a non-zero  $p$  such that for all  $x \in C$  it holds that*

$$\langle p, x \rangle \leq \langle p, x_0 \rangle.$$

*Proof.* For every  $x \in C$  we define

$$F_x := \{p \mid \|p\|_2 = 1, \langle p, x \rangle \leq \langle p, x_0 \rangle\}$$

and observe that each  $F_x$  is closed and since  $F_x$  is subset of the compact set  $\{\|p\|_2 = 1\}$ , it is compact as well. Now we show that every intersection of finitely many  $F_x$  is non-empty: For  $x_1, \dots, x_n \in C$  we define

$$M := \text{conv}(x_1, \dots, x_n).$$

Since  $C$  is convex, we have  $M \subset C$  and hence,  $x_0 \notin M$ . Since  $M$  is non-empty, convex and closed, we can invoke the Projection Theorem to get that for all  $x \in M$  it holds that

$$\langle x_0 - P_M x_0, x - P_M x_0 \rangle \leq 0$$

from which we deduce

$$\|x_0 - P_M x_0\|_2^2 \leq \langle x_0 - P_M x_0, x_0 - x \rangle.$$

Now we set  $p = (x_0 - P_M x_0) / \|x_0 - P_M x_0\|$  and get  $\langle p, x \rangle \leq \langle p, x_0 \rangle$  for all  $x \in M$  and especially  $\langle p, x_i \rangle \leq \langle p, x_0 \rangle$  for  $i = 1, \dots, n$ . This shows  $p \in \bigcap_{i=1}^n F_{x_i}$ . By the previous proposition, we conclude that  $\bigcap_{x \in C} F_x$  is non-empty as well, which shows the assertion.  $\square$

**Theorem 3.8** (Separation Theorem). *Any two non-empty closed, convex and disjoint sets  $C_1$  and  $C_2$  can be separated by a hyperplane.*

*Proof.* We consider the Minkowski sum  $C = C_1 + (-C_2)$  which is non-empty, closed and convex with  $0 \notin C$ . By Proposition 3.7 we can separate 0 from  $C$  and this hyperplane also separates  $C_1$  and  $C_2$ .  $\square$

We will also use the following slight generalization which even characterizes proper separation:

**Theorem 3.9** (Proper separation theorem). *Two non-empty, convex sets  $C_1$  and  $C_2$  can be properly separated if and only if  $\text{ri } C_1$  and  $\text{ri } C_2$  are disjoint.*

*Proof.* “ $\Rightarrow$ ”: Suppose that  $H_{p,\alpha}$  separates  $C_1$  and  $C_2$  properly, i.e. for all  $x_i \in C_i$  we have  $\langle p, x_1 \rangle \leq \alpha \leq \langle p, x_2 \rangle$  and there exist  $\bar{x}_i \in C_i$  such that  $\langle p, \bar{x}_1 \rangle < \langle p, \bar{x}_2 \rangle$ . We aim to show that for all  $x_i \in \text{ri } C_i$  we have that  $\langle p, x_1 \rangle < \langle p, x_2 \rangle$  as well (which then implies that  $\text{ri } C_1$  and  $\text{ri } C_2$  are disjoint).

To do so, assume that there are  $x_i \in \text{ri } C_i$  such that  $\langle p, x_1 \rangle = \langle p, x_2 \rangle$ . By Proposition 1.8 we know that there exists  $\epsilon > 0$  such that for  $i = 1, 2$  we have

$$y_i := x_i - \epsilon(\bar{x}_i - x_i) = (1 + \epsilon)x_i - \epsilon\bar{x}_i \in C_i.$$

Thus,

$$\begin{aligned} \langle p, y_1 \rangle &= \langle p, (1 + \epsilon)x_1 - \epsilon\bar{x}_1 \rangle \\ &= \langle p, (1 + \epsilon)x_2 \rangle - \epsilon\langle p, \bar{x}_1 \rangle \\ &> \langle p, (1 + \epsilon)x_2 \rangle - \epsilon\langle p, \bar{x}_2 \rangle \\ &= \langle p, y_2 \rangle. \end{aligned}$$

This contradicts the separation property of  $H_{p,\alpha}$ .

“ $\Leftarrow$ ”: Now let  $\text{ri}(C_1) \cap \text{ri}(C_2) = \emptyset$ , i.e.  $0 \notin \text{ri}(C_1) - \text{ri}(C_2) = \text{ri}(C_1 - C_2)$ . We construct a properly separating hyperplane  $H_{p,\alpha}$ .

Remembering the remark after Proposition 1.8, we assume without loss of generality, that  $\text{aff}(C_1 - C_2) = \mathbb{R}^d$ . Then  $\text{ri}(C_1 - C_2) = \text{int}(C_1 - C_2)$  and by Proposition 3.7 there exists  $p \neq 0$  such that for all  $x_1 - x_2 \in \text{int}(C_1 - C_2)$

$$\langle p, x_1 - x_2 \rangle \geq 0.$$

We conclude that

$$\{x \in \mathbb{R}^d \mid \langle p, x \rangle \geq 0\} \supset \text{int}(C_1 - C_2).$$

Since the set on the left is closed, we also get

$$\{x \in \mathbb{R}^d \mid \langle p, x \rangle \geq 0\} \supset \text{cl}(C_1 - C_2).$$

Thus, we even have

$$\langle p, x_1 - x_2 \rangle \geq 0$$

for all  $x_1, x_2$  with  $x_1 - x_2 \in C_1 - C_2$ . Thus, we can separate  $C_1$  and  $C_2$  with this  $p$  and  $\alpha := \sup_{x_2 \in C_2} \langle p, x_2 \rangle$ . Finally, for  $x_1 - x_2 \in \text{int}(C_1 - C_2)$  we have the strict inequality  $\langle p, x_1 - x_2 \rangle > 0$ .

□

For a non-empty convex set  $C$  one calls a halfspace *supporting*, if it contains  $C$  and has a point in the closure of  $C$  in its boundary. A *supporting hyperplane*, is one which is the boundary of a supporting halfspace. Our above results show that for every  $x_0 \in \text{cl}(C) \setminus \text{int}(C)$ , there exists a supporting hyperplane for  $C$  which contains  $x_0$  (namely, one which separates  $x_0$  from  $\text{int}(C)$ ).

## 4 Convex functions

When working with convex functions, it will be convenient to use the extended real numbers  $\bar{\mathbb{R}} = [-\infty, \infty]$ . We will use the following conventions:

$$\begin{aligned} \text{For } -\infty < a \leq \infty: & \quad a + \infty = \infty + a = \infty, \\ \text{For } -\infty \leq a < \infty: & \quad a - \infty = -\infty + a = -\infty, \\ \text{For } 0 < a \leq \infty: & \quad a\infty = \infty a = \infty, \quad a(-\infty) = (-\infty)a = -\infty, \\ \text{For } -\infty \leq a < 0: & \quad a\infty = \infty a = -\infty, \quad a(-\infty) = (-\infty)a = \infty, \\ & \quad 0\infty = \infty 0 = 0 = 0(-\infty) = (-\infty)0, \\ & \quad -(-\infty) = \infty, \\ & \quad \inf \emptyset = \infty, \\ & \quad \sup \emptyset = -\infty. \end{aligned}$$

The expressions  $\infty - \infty$  and  $-\infty + \infty$  will be left undefined intentionally.

We will use the notion of *indicator function* of a set  $S$  which is  $i_S : \mathbb{R}^d \rightarrow \bar{\mathbb{R}}$  defined by

$$i_S(x) := \begin{cases} 0, & \text{if } x \in S \\ \infty, & \text{if } x \notin S. \end{cases}$$

**Definition 4.1.** A function  $f : \mathbb{R}^d \rightarrow \bar{\mathbb{R}}$  is proper, if  $f(x) > -\infty$  for all  $x$  and there exists  $x_0$  such that  $f(x_0) < \infty$ . The (*effective*) *domain* of  $f$  is

$$\text{dom } f := \{x \mid f(x) < \infty\}.$$

We say that  $f$  is *lower semicontinuous* (lsc) at  $x_0$  if

$$f(x_0) \leq \liminf_{x \rightarrow x_0} f(x)$$

and  $f$  is called lsc if it is lsc at every point.

It is a simple observation that an indicator function  $i_S$  is lower semicontinuous if and only if  $S$  is closed.

**Example 4.2.** Consider  $f_i : \mathbb{R} \rightarrow \bar{\mathbb{R}}$  given by

$$\begin{aligned} f_1 &= i_{[1,1]}, & f_2 &= i_{[-1,1]} \\ f_3(x) &= \begin{cases} 0, & \text{if } x = 0 \\ 1, & \text{else.} \end{cases} & f_4(x) &= \begin{cases} \frac{1}{x}, & \text{if } x > 0 \\ \infty, & \text{else.} \end{cases} \end{aligned}$$

Then:  $f_1$  is not lsc while  $f_2, f_3$  and  $f_4$  are. △

**Definition 4.3.** The *epigraph* of a function  $f : \mathbb{R}^d \rightarrow \bar{\mathbb{R}}$  is

$$\text{epi } f := \{(x, \alpha) \in \mathbb{R}^d \times \mathbb{R} \mid f(x) \leq \alpha\}.$$

The *level sets* of  $f$  for level  $\alpha \in \mathbb{R}$  are

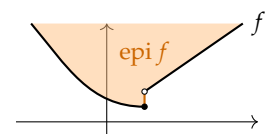
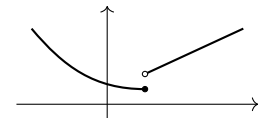
$$\text{lev}_\alpha f := \{x \in \mathbb{R}^d \mid f(x) \leq \alpha\}$$

Indicator functions are helpful to formulate constrained minimization problems  $\min_{x \in S} f(x)$  as (formally) unconstrained minimization problems  $\min_{x \in \mathbb{R}^d} f(x) + i_S(x)$ .

Equivalent formulation of lower semicontinuity are

- $f$  is lsc at  $x_0$  if for every  $\lambda$  with  $\lambda < f(x_0)$  there exists  $\epsilon > 0$  such that  $f(x) \geq \lambda$  for all  $x \in B_\epsilon(x_0)$ .
- $f$  is lsc at  $x_0$  if  $f(x_0) \leq \lim_{k \rightarrow \infty} f(x_k)$  whenever  $x_k \rightarrow x_0$  and  $\lim f(x_k)$  exists.

Intuitively, a function is lsc, if it only may jump down in the limit.





**Theorem 4.4.** *The following conditions for  $f : \mathbb{R}^d \rightarrow \mathbb{R}$  are equivalent:*

- i)  $f$  is lsc,
- ii)  $\text{epi } f$  is closed,
- iii) for all  $\alpha \in \mathbb{R}$ ,  $\text{lev}_\alpha f$  is closed.

**Proof.** **i)  $\implies$  ii):** Let  $f$  be lsc and consider a sequence  $(x_n, \alpha_n) \rightarrow (x, \alpha)$  with  $(x_n, \alpha_n) \in \text{epi } f$ , i.e.  $f(x_n) \leq \alpha_n$ . Since  $f$  is lsc, we have  $f(x) \leq \liminf f(x_n) \leq \lim \alpha_n = \alpha$  and this shows  $(x, \alpha) \in \text{epi } f$ .

**ii)  $\implies$  iii):** Let  $\text{epi } f$  be closed and  $\alpha \in \mathbb{R}$ . If  $x_n$  converges to  $x$  and fulfills  $x_n \in \text{lev}_\alpha f$ , we also have  $\lim(x_n, \alpha) = (x, \alpha)$  and  $(x_n, \alpha) \in \text{epi } f$ . Since  $\text{epi } f$  is closed, we have  $(x, \alpha) \in \text{epi } f$  which shows  $f(x) \leq \alpha$ , and hence,  $x \in \text{lev}_\alpha f$  which shows the closedness.

**iii)  $\implies$  i):** Let  $\text{lev}_\alpha f$  be closed for all  $\alpha$  and let  $x_0 \in \mathbb{R}^d$ . If  $f(x_0) = -\infty$  holds,  $f$  is obviously lsc at  $x_0$ . Otherwise, let  $\alpha < f(x_0)$  which means that  $x_0 \notin \text{lev}_\alpha f$ . Since  $\text{lev}_\alpha f$  is closed, there is  $\epsilon > 0$  such that  $\text{lev}_\alpha f \cap B_\epsilon(x_0) = \emptyset$ . Thus,  $f(x) > \alpha$  for all  $x \in B_\epsilon(x_0)$  and hence,  $f$  is lsc at  $x_0$ . □

**Proposition 4.5.** *If  $f, g, f_i : \mathbb{R}^d \rightarrow \bar{\mathbb{R}}$  are lsc,  $A \in \mathbb{R}^{d \times m}$  and  $\alpha > 0$ , then the following functions are also lsc:*

- i)  $f + g$  (if it's defined),  $\alpha f$
- ii)  $f \circ A$
- iii)  $\inf(f, g)$
- iv)  $\sup_i f_i$  for arbitrarily many  $f_i$ .

**Proof.** The items i) and ii) follow directly from the definition. For iii) note that  $(x, \alpha) \in \text{epi}(\inf(f, g))$  if and only if  $f(x) \leq \alpha$  or  $g(x) \leq \alpha$ . Hence,  $\text{epi}(\inf(f, g)) = \text{epi } f \cup \text{epi } g$ . And since the union of two closed sets is closed, the assertion follows.

For item iv) note that  $\text{epi}(\sup_i f_i)$  is the intersection of all the sets  $\text{epi}(f_i)$  and since the intersection of closed sets is closed set, we are done. □

Since we characterized lower semi-continuity of a function by closedness of the epigraph, we can associate to every function  $f : \mathbb{R}^d \rightarrow \bar{\mathbb{R}}$  which is not a lsc another lsc function  $\bar{f}$  via closing the epigraph, i.e.  $\bar{f}$  is characterized by

$$\text{epi}(\bar{f}) = \overline{\text{epi}(f)}.$$

This function is called the *lower semi-continuous hull* of  $f$ .

It's not true that the infimum of infinitely many lsc  $f_i$  is again lsc. Can you think of an example?

Taking special care of the case when  $f(x) = -\infty$  can occur, we define the *closure* of  $f$  by

$$\text{cl}(f) := \begin{cases} \bar{f}, & \text{if } f(x) > -\infty \text{ for all } x \\ -\infty, & \text{if } f(x) = -\infty \text{ for some } x. \end{cases}$$

We call a function closed, if  $\text{cl } f = f$  and for functions which do not take the value  $-\infty$ , closed means the same as lsc and hence, we have

$$(\text{cl } f)(x_0) = \liminf_{x \rightarrow x_0} (\text{cl } f)(x) = \liminf_{x \rightarrow x_0} f(x) \leq f(x_0).$$

Intuitively, the closure of a function moves the function values to the lowest possible values at points of discontinuity.

*Example 4.6.* For the function

$$f(x) = i_{]a,b[}(x) = \begin{cases} 0, & a < x < b \\ \infty, & \text{else,} \end{cases}$$

the closure (and also the lsc hull) is

$$\text{cl } f = \bar{f} = i_{[a,b]}.$$

△

**Definition 4.7.** A function  $f : \mathbb{R}^d \rightarrow \bar{\mathbb{R}}$  is called

- *convex*, if for all  $x, y \in \text{dom}(f)$ ,  $\lambda \in [0, 1]$  it holds that

$$f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda)f(y).$$

- *strictly convex*, if it is convex and for  $x \neq y$  and  $\lambda \in ]0, 1[$  it holds that

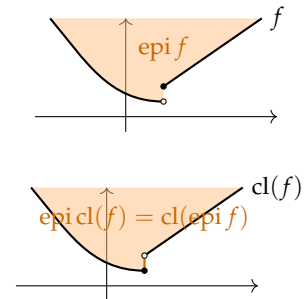
$$f(\lambda x + (1 - \lambda)y) < \lambda f(x) + (1 - \lambda)f(y).$$

- *uniformly convex*, if there exists a strictly increasing function  $\varphi$  with  $\varphi(0) = 0$  such that for all  $x, y \in \text{dom}(f)$ ,  $\lambda \in [0, 1]$  it holds that

$$f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda)f(y) - \lambda(1 - \lambda)\varphi(\|x - y\|_2).$$

- *strongly convex* with constant  $\sigma > 0$ , if for all  $x, y \in \text{dom}(f)$ ,  $\lambda \in [0, 1]$  it holds that

$$f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda)f(y) - \frac{\sigma}{2}\lambda(1 - \lambda)\|x - y\|_2^2.$$



One can show that a function  $f$  is strongly convex with constant  $\sigma$  if and only if the function  $x \mapsto f(x) - \frac{\sigma}{2} \|x\|_2^2$  is convex.

It's clear that

$$\text{strongly convex} \implies \text{strictly convex} \implies \text{convex}$$

but the reverse implications do not hold:  $f(x) = i_{[a,b]}$  is convex but not strictly so,  $f(x) = x^4$  is strictly convex, but not strongly so. It's also clear that  $\text{dom}(f)$  is a convex set for any convex function  $f$ .

Finally, we call a function  $f$  *concave* if  $-f$  is convex.

By induction one can deduce from the defining in the inequality:

**Lemma 4.8** (Jensen's inequality). *A function  $f : \mathbb{R}^d \rightarrow \bar{\mathbb{R}}$  is convex if and only if for all  $x_i \in \text{dom}(f)$  and  $\lambda_i \in [0, 1]$  with  $\sum_{i=1}^n \lambda_i = 1$  it holds that*

$$f\left(\sum_{i=1}^n \lambda_i x_i\right) \leq \sum_{i=1}^n \lambda_i f(x_i).$$

**Lemma 4.9.** *If a convex function  $f$  has a finite value at some point  $x_0 \in \text{ri}(\text{dom}(f))$ , then it is proper (i.e. it does not assume the value  $-\infty$  anywhere).*

*Proof.* For a contradiction, assume that  $x_1 \in \text{dom}(f)$  exists with  $f(x_1) = -\infty$ . Since  $x_0 \in \text{ri}(\text{dom}(f))$ , there exists  $x_2 \in \text{ri}(\text{dom}(f))$  and  $\lambda \in [0, 1]$  such that  $x_0 = \lambda x_1 + (1 - \lambda)x_2$ . But, by convexity we would get

$$f(x_0) \leq \lambda f(x_1) + (1 - \lambda)f(x_2) = -\infty$$

which contradicts  $f(x_0)$  finite.  $\square$

**Example 4.10.** 1. Any affine function is convex as well as concave and only affine functions are both at the same time.

2. Any norm  $f(x) = \|x\|$  is convex, but no norm is strictly convex since for  $y = 0$  and  $x \neq 0$  and  $\lambda \in [0, 1]$  we have

$$\|\lambda x + (1 - \lambda)y\| = \lambda \|x\| + (1 - \lambda)\|y\|.$$

3. Indicator functions  $i_C$  are convex exactly for convex sets  $C$ .

$\triangle$

Put differently: A function is strongly convex (with modulus)  $\sigma$ , if it is uniformly convex with respect to  $\varphi(t) = \frac{\sigma}{2} t^2$ .

The constant of strong convexity is also called *modulus of strong convexity*.

## 5 Characterization of convex functions

For differentiable functions, convexity can be described with derivatives:

**Theorem 5.1.** *If  $f : \mathbb{R}^d \rightarrow \mathbb{R}$  is differentiable, the following conditions are equivalent:*

i)  $f$  is convex,

ii) the gradient  $\nabla f : \mathbb{R}^d \rightarrow \mathbb{R}^d$  is monotone, i.e. for all  $x, y \in \mathbb{R}^d$

$$\langle \nabla f(x) - \nabla f(y), x - y \rangle \geq 0,$$

iii) for all  $x, y \in \mathbb{R}^d$  it holds that

$$f(y) \geq f(x) + \langle \nabla f(x), y - x \rangle.$$

If  $f$  is twice differentiable, then  $f$  is convex if and only if the Hessian  $\nabla^2 f(x)$  is positive semi-definite for every  $x$ .

Strict convexity is characterized by strict inequalities for  $x \neq y$  in i) and ii). Positive definiteness in the case of twice differentiability is sufficient but not necessary for strict convexity.

*Proof.* We first show  $i) \iff iii) \iff ii)$ :

**i)  $\implies$  iii):** For  $\lambda \in ]0, 1]$  we rearrange  $f(\lambda y + (1 - \lambda)x) \leq \lambda f(y) + (1 - \lambda)f(x)$  to

$$\frac{f(x + \lambda(y - x)) - f(x)}{\lambda} \leq f(y) - f(x).$$

For  $\lambda \rightarrow 0$  the left hand side converges to the directional derivative of  $f$  in  $x$  in the direction of  $y - x$  which equals  $\langle \nabla f(x), y - x \rangle$  since  $f$  is differentiable.

**iii)  $\implies$  i):** Set  $x_\lambda = \lambda x + (1 - \lambda)y$  and note that  $x - x_\lambda = (1 - \lambda)(x - y)$  and  $y - x_\lambda = -\lambda(x - y)$ . We get the inequalities

$$f(x) \geq f(x_\lambda) + \langle \nabla f(x_\lambda), x - x_\lambda \rangle = f(x_\lambda) + (1 - \lambda)\langle \nabla f(x_\lambda), x - y \rangle$$

$$f(y) \geq f(x_\lambda) + \langle \nabla f(x_\lambda), y - x_\lambda \rangle = f(x_\lambda) - \lambda\langle \nabla f(x_\lambda), x - y \rangle.$$

Multiplying the first inequality with  $\lambda$  and the second with  $(1 - \lambda)$  and adding the inequality shows the assertion.

**iii)  $\implies$  ii):** We add  $f(y) \geq f(x) + \langle \nabla f(x), y - x \rangle$  and  $f(x) \geq f(y) + \langle \nabla f(y), x - y \rangle$  and get  $0 \geq \langle \nabla f(x) - \nabla f(y), y - x \rangle$  which shows the assertion.

A function  $g : \mathbb{R}^d \rightarrow \mathbb{R}^d$  is called *monotone* if for all  $x, y$  it holds that  $\langle g(x) - g(y), x - y \rangle \geq 0$ . Can you see, why this is called “monotonicity”? (Hint: consider  $d = 1$ .)

The equivalence of i) and iii) is quite remarkable: Besides the definition of convex function from above by “function lies below its secants” we have the equivalent description from below by “function lies above its tangents”. This is comparable to the duality of the descriptions of convex sets from the inside (via convex combinations) and the outside (by separating hyperplanes).

The function  $f(x) = x^4$  is strictly convex, but the second derivative vanishes (i.e., is not positive definite) at  $x = 0$ .

ii)  $\implies$  iii): Since for  $h(\lambda) = f(y + \lambda(x - y))$  we have  $h'(\lambda) = \langle \nabla f(y + \lambda(x - y)), x - y \rangle$  and the fundamental theorem of calculus gives

$$\begin{aligned} f(x) - f(y) &= h(1) - h(0) = \int_0^1 h'(\lambda) d\lambda \\ &= \int_0^1 \langle \nabla f(y + \lambda(x - y)), x - y \rangle d\lambda. \end{aligned}$$

Subtracting  $\langle \nabla f(x), x - y \rangle$  from both sides, we get

$$f(x) - f(y) - \langle \nabla f(x), x - y \rangle = \int_0^1 \langle \nabla f(y + \lambda(x - y)) - \nabla f(x), x - y \rangle d\lambda.$$

The right hand side is non-negative (since  $y + \lambda(x - y) - x = (1 - \lambda)(y - x)$  and  $\nabla f$  is monotone), by assumption and this shows i).

The claim on strict convexity follows by inspection of the above arguments in this case.

For twice differentiable functions we show the equivalence of positive semidefinite Hessian and ii): Set  $x_\tau = x + \tau s$  for  $\tau > 0$  and with ii) we get (using  $\frac{d}{d\lambda} \langle \nabla f(x + \lambda s), s \rangle = \langle \nabla^2 f(x + \lambda s) s, s \rangle$ )

$$\begin{aligned} 0 &\leq \frac{1}{\tau^2} \langle \nabla f(x_\tau) - f(x), x_\tau - x \rangle \\ &= \frac{1}{\tau} \langle \nabla f(x_\tau) - f(x), s \rangle = \frac{1}{\tau} \int_0^\tau \langle \nabla^2 f(x + \lambda s) s, s \rangle d\lambda \end{aligned}$$

and  $\tau \rightarrow 0$  shows that  $\nabla^2 f(x) \succeq 0$ . Conversely, if  $\nabla^2 f(x) \succeq 0$  we can write (using the fundamental theorem of calculus twice)

$$\begin{aligned} f(y) &= f(x) + \langle \nabla f(x), y - x \rangle + \int_0^1 \int_0^\tau \langle \nabla^2 f(x + \lambda(y - x))(y - x), y - x \rangle d\lambda d\tau \\ &\geq f(x) + \langle \nabla f(x), y - x \rangle \end{aligned}$$

which implies convexity of  $f$ .

□

**Example 5.2.** 1. The function  $f(x) = \frac{1}{2} \langle Ax, x \rangle + \langle a, x \rangle + b$  is convex as soon as the matrix  $A$  is positive semidefinite and it is strictly convex if  $A$  is positive definite (since  $\nabla^2 f(x) = A$ ). Moreover, one can even show strong convexity of  $f$  for positive definite  $A$  (what is the modulus?).

2. The “log-sum-exp” function is

$$\text{logsumexp}(x) = \log\left(\sum_{i=1}^d \exp(x_i)\right).$$

We abbreviate  $h(x) = \sum_{i=1}^d \exp(x_i)$  and get

$$\partial_i \text{logsumexp}(x) = \frac{\exp(x_i)}{h(x)}$$

and

$$\partial_j \partial_i \log \text{sumexp}(x) = \begin{cases} \frac{\exp(x_i) h(x) - \exp(2x_i)}{h(x)^2}, & i = j \\ -\frac{\exp(x_i + x_j)}{h(x)^2}, & i \neq j. \end{cases}$$

Thus, we can compute for  $z \in \mathbb{R}^d$ :

$$\begin{aligned} \langle \nabla^2 \log \text{sumexp}(x) z, z \rangle &= \frac{1}{h(x)^2} \left( h(x) \sum_{i=1}^d \exp(x_i) z_i^2 - \sum_{i,j=1}^d \exp(x_i + x_j) z_i z_j \right) \\ &= \frac{1}{h(x)^2} \left( \sum_{i,j=1}^d \exp(x_i + x_j) z_i^2 - \sum_{i,j=1}^d \exp(x_i + x_j) z_i z_j \right) \\ &= \frac{1}{h(x)^2} \left( \frac{1}{2} \sum_{i,j=1}^d \exp(x_i + x_j) z_i^2 + \frac{1}{2} \sum_{i,j=1}^d \exp(x_i + x_j) z_j^2 \right. \\ &\quad \left. - \sum_{i,j=1}^d \exp(x_i + x_j) z_i z_j \right) \\ &= \frac{1}{h(x)^2} \left( \sum_{i,j=1}^d \exp(x_i + x_j) \underbrace{\left( \frac{1}{2} z_i^2 + \frac{1}{2} z_j^2 - z_i z_j \right)}_{=\frac{1}{2}(z_i - z_j)^2 \geq 0} \right) \geq 0 \end{aligned}$$

which shows convexity of  $\log \text{sumexp}$ .

△

If a function  $f$  is convex, its level set  $\text{lev}_\alpha f$  are all convex (if  $f(x), f(y) \leq \alpha$ , then  $f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda)f(y) \leq \alpha$ ). The converse is not true (consider something like  $f(x) = \log(1+x^2)$  on the real line, for example). Convexity of the epigraph, though, does characterize convexity of the function:

Functions with convex level sets sometimes are called *quasi-convex*.

**Proposition 5.3.** A function  $f : \mathbb{R}^d \rightarrow \bar{\mathbb{R}}$  is convex if and only if  $\text{epi } f$  is a convex set.

*Proof.* The case  $f \equiv \infty$  is clear since in this case  $\text{epi } f = \emptyset$ . So consider  $\text{dom}(f) \neq \emptyset$ .

⇒: Let  $f$  be convex and  $(x, a), (y, b) \in \text{epi}(f)$ , i.e.  $f(x) \leq a$ ,  $f(y) \leq b$ . Hence, for  $\lambda \in [0, 1]$ :  $f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda)f(y) \leq \lambda a + (1-\lambda)b$ . But this means that  $\lambda \begin{bmatrix} x \\ a \end{bmatrix} + (1-\lambda) \begin{bmatrix} y \\ b \end{bmatrix} \in \text{epi}(f)$ .

⇐: Let  $\text{epi}(f)$  be convex and  $x, y \in \text{dom}(f)$  with  $f(x), f(y) \neq -\infty$ . Since  $(x, f(x)), (y, f(y)) \in \text{epi}(f)$ , we have for  $\lambda \in [0, 1]$

Beware: Sometimes I write tuples as  $(x, a)$  and sometimes they will be  $\begin{bmatrix} x \\ a \end{bmatrix}$  depending on the typographic circumstances.

$$\lambda(x, f(x)) + (1-\lambda)(y, f(y)) = (\lambda x + (1-\lambda)y, \lambda f(x) + (1-\lambda)f(y)) \in \text{epi}(f)$$

and this means that  $f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda)f(y)$ . If  $f(x) = -\infty$ , then we use  $(x, -N)$  instead of  $(x, f(x))$  and let  $N \rightarrow -\infty$ . □

We apply the separation of points from convex sets with hyperplanes to the epigraph and obtain:

**Proposition 5.4.** If  $f : \mathbb{R}^d \rightarrow \bar{\mathbb{R}}$  is a convex and proper, then there exists  $p \in \mathbb{R}^d$  and  $\alpha \in \mathbb{R}$  such that for all  $x$  it holds that

$$f(x) \geq \langle p, x \rangle + \alpha.$$

*Proof.* Since  $\text{cl } f \leq f$  we only need to consider the case where  $f$  is closed.

If  $x \notin \text{dom}(f)$ , the inequality is valid for any  $p$  and  $\alpha$ . Now fix  $x_0 \in \text{dom}(f)$  and  $\beta$  such that  $f(x_0) > \beta$ , i.e.  $(x_0, \beta) \notin \text{epi}(f)$ . Since  $\text{epi}(f)$  is closed and convex, we can use Theorem 3.5 to separate the compact singleton  $\{(x_0, \beta)\}$  from  $\text{epi}(f)$ . This means, that there exists  $(\bar{p}, -b) \in \mathbb{R}^{d+1} \setminus \{0\}$  and  $\epsilon > 0$  such that for all  $x \in \text{dom}(f)$  it holds that

$$\langle \bar{p}, x \rangle - bf(x) \leq \langle \bar{p}, x_0 \rangle - b\beta - \epsilon. \quad (*)$$

For  $x = x_0$  we get

$$b(f(x_0) - \beta) \geq \epsilon > 0$$

and since  $f(x_0) - \beta > 0$  we obtain  $b > 0$  as well. Hence, with  $p = \bar{p}/b$  we get, by dividing  $(*)$  by  $b$ ,

$$f(x) \geq \langle p, x \rangle - \langle p, x_0 \rangle + \beta$$

and we proved the claim with  $\alpha = -\langle p, x_0 \rangle + \beta$ . □

The following corollary is a simple exercise.

**Corollary 5.5.** If  $f : \mathbb{R}^d \rightarrow \bar{\mathbb{R}}$  is convex, it holds for all  $x_0 \in \mathbb{R}^d$

$$\text{cl } f(x_0) = \sup\{\langle p, x_0 \rangle + \alpha \mid p \in \mathbb{R}^d, \alpha \in \mathbb{R} \text{ with } f(x) \geq \langle p, x \rangle + \alpha \forall x \in \mathbb{R}^d\}.$$

Convexity is preserved under several operations.

**Proposition 5.6.** 1.  $f_1 + f_2$  is convex if  $f_1$  and  $f_2$  are convex.

2.  $\alpha f$  is convex if  $f$  is convex and  $\alpha \geq 0$ .

3.  $f \circ A$  is convex if  $f$  is convex and  $A$  is linear.

4.  $\varphi \circ f$  is convex if  $f$  is convex and  $\varphi : \bar{\mathbb{R}} \rightarrow \bar{\mathbb{R}}$  is convex and increasing (with  $\varphi(\infty) = \infty$  and  $\varphi(-\infty) = -\infty$ ).

5.  $\sup_{i \in I} f_i$  is convex if all the  $f_i$  are.

All proofs are straightforward calculations.

**Example 5.7.** By Proposition 5.6 5. we see that  $f(x) = \max(x_1, \dots, x_d)$  is convex. △

## 6 Continuity of convex functions and minimizers

Surprisingly, the notion of convexity implies a certain continuity. We start with a lemma:

**Lemma 6.1.** *Let  $f : \mathbb{R}^d \rightarrow \bar{\mathbb{R}}$  be proper and convex and suppose that it is locally bounded at  $x_0$ , i.e. there exists  $\delta > 0$  and  $m, M \in \mathbb{R}$  such that for  $x \in B_{2\delta}(x_0)$  it holds that*

$$m \leq f(x) \leq M.$$

*Then  $f$  is Lipschitz continuous on  $B_\delta(x_0)$  with Lipschitz constant at most  $(M - m)/\delta$ .*

*Proof.* The proof is a bit technical: Let  $x_1, x_2 \in B_\delta(x_0)$ ,  $x_1 \neq x_2$  and set

$$y := \left(1 + \frac{\delta}{\|x_1 - x_2\|_2}\right)x_2 - \frac{\delta}{\|x_1 - x_2\|_2}x_1 = x_2 + \delta \frac{x_1 - x_2}{\|x_1 - x_2\|_2^2}.$$

It holds that  $\|y - x_2\|_2 \leq \delta$ , i.e.  $y \in B_{2\delta}(x_0)$ . By rearranging to

$$x_2 = \frac{\|x_1 - x_2\|_2}{\delta + \|x_1 - x_2\|_2}y + \frac{\delta}{\delta + \|x_1 - x_2\|_2}x_1$$

we see that  $x_2$  is a convex combination of  $x_1$  and  $y$ .

Since  $f$  is convex we get,

$$f(x_2) \leq \frac{\|x_1 - x_2\|_2}{\delta + \|x_1 - x_2\|_2}f(y) + \frac{\delta}{\delta + \|x_1 - x_2\|_2}f(x_1)$$

which leads to

$$\begin{aligned} f(x_2) - f(x_1) &\leq \frac{\|x_1 - x_2\|_2}{\delta + \|x_1 - x_2\|_2}(f(y) - f(x_1)) \\ &\leq \frac{M - m}{\delta} \|x_1 - x_2\|_2. \end{aligned}$$

Since we can swap the roles of  $x_1$  and  $x_2$  this shows the desired Lipschitz continuity.  $\square$

**Theorem 6.2.** *Let  $f : \mathbb{R}^d \rightarrow \bar{\mathbb{R}}$  be proper and convex. Then  $f$  is Lipschitz continuous on any compact subset  $C$  of  $\text{ri}(\text{dom}(f))$ .*

*Proof.* 1. (Restriction to fulldimensional case.) Let  $C \subset \text{ri}(\text{dom}(f))$ .

Since  $C$  is a subset of the affine hull of the relative interior, we can restrict  $f$  to this affine set and hence, we may assume that  $\text{ri}(\text{dom}(f)) = \text{int}(\text{dom}(f))$ .

2. (Locally Lipschitz.) Let  $x_0 \in C$  and since  $x_0 \in \text{int}(\text{dom}(f))$ , there exist  $v_1, \dots, v_r \in \text{dom}(f)$  and  $\delta > 0$  such that

$$B_{2\delta}(x_0) \subset S := \text{conv}(v_1, \dots, v_r) \subset \text{dom}(f).$$

Hence, any  $v \in B_{2\delta}(x_0)$  can be written as  $v = \sum_{i=1}^r \alpha_i v_i$ ,  $\sum_i \alpha_i = 1$ ,  $\alpha_i \geq 0$  and convexity of  $f$  gives

$$f(v) \leq \sum_{i=1}^r \alpha_i f(v_i) \leq \max_i f(v_i) =: M.$$



This shows that  $f$  is bounded from above on  $B_{2\delta}(x_0)$ .

By Proposition 5.4 we know that there are  $p \in \mathbb{R}^d$  and  $\alpha \in \mathbb{R}$  such that

$$f(x) \geq \langle p, x \rangle + \alpha.$$

Hence,  $f$  is bounded from below on  $B_{2\delta}(x_0)$  since the right hand side is so.

Thus, we can apply Lemma 6.1 and obtain that  $f$  is Lipschitz continuous on  $B_\delta(x_0)$ .

3. (Lipschitz on full  $C$ ) The previous step shows that for any  $x_0 \in C$  there exists  $\delta(x_0) > 0$  and  $L(x_0, \delta) \geq 0$  such that for all  $x, y \in B_\delta(x_0)$  it holds that

$$|f(x) - f(y)| \leq L(x_0, \delta) \|x - y\|_2.$$

The balls  $B_{\delta(x_0)}(x_0)$  cover the compact set  $C$  and hence, there are  $x_1, \dots, x_K$  such that  $C$  is covered by  $B_{\delta(x_1)}(x_1), \dots, B_{\delta(x_K)}(x_K)$ . Hence  $f$  is Lipschitz continuous on  $C$  with constant  $L = \max_k (L(x_k, \delta(x_k)))$ .

□

**Corollary 6.3.** *If  $f : \mathbb{R}^d \rightarrow \bar{\mathbb{R}}$  is proper and convex, then it is continuous on  $\text{ri}(\text{dom}(f))$  relative to  $\text{aff}(\text{dom}(f))$ . If  $f$  is additionally finite everywhere, then it is continuous on  $\mathbb{R}^d$ .*

Now we start our treatment of minimization problems

$$\min_{x \in \mathbb{R}^d} f(x).$$

A ridiculously large class of practically relevant problems can be written in this form (note that we can treat constraint problems just by setting  $f = g + i_C$ ) and we will see some examples later in the lecture.

In addition to the usual definitions of local and global minima of functions, we will also need the set of minimizers which we will denote by

$$\text{argmin } f := \{\hat{x} \in \text{dom}(f) \mid f(\hat{x}) = \inf_x f(x)\}.$$

Similarly we have  $\text{argmax } f := \text{argmin}(-f)$ .

Now we are interested in conditions on  $f$  which ensure the following properties

- i)  $f$  attains its minimum, i.e.  $\text{argmin } f \neq \emptyset$ ,
- ii) every local minimizer of  $f$  is a global minimizer,
- iii) the global minimizer is unique.

Mere convexity does not even ensure existence of minimizers, even in the proper case as  $f(x) = \exp(x)$  shows.

**Definition 6.4.** A function  $f : \mathbb{R}^d \rightarrow \bar{\mathbb{R}}$  is called *level-bounded* or *coercive* if all level sets  $\text{lev}_\alpha f$  are bounded.

Another way to put this:  $f$  is coercive if and only if  $f(x) \rightarrow \infty$  whenever  $\|x\|_2 \rightarrow \infty$ .

$\Leftarrow$ : Assume that  $f(x) \rightarrow \infty$  whenever  $\|x\|_2 \rightarrow \infty$ . If  $x \in \text{lev}_\alpha f$ , then  $f(x) \leq \alpha$ . Thus, there can't be an unbounded sequence in  $\text{lev}_\alpha f$  since this would contradict that  $f(x) \leq \alpha$  for all  $x \in \text{lev}_\alpha f$ .

$\Rightarrow$ : We prove the contraposition: Assume that there exists  $\|x_n\| \rightarrow \infty$  with  $f(x_n) \leq C$ . But then  $\text{lev}_C f$  is not bounded, since  $x_n \in \text{lev}_C f$ .

**Theorem 6.5.** Let  $f : \mathbb{R}^d \rightarrow \bar{\mathbb{R}}$  be proper.

- i) If  $f$  is lsc and coercive, then  $\text{argmin } f$  is non-empty and compact.
- ii) If  $f$  is convex, then every local minimizer is also global and the set  $\text{argmin } f$  is convex (but possibly empty).
- iii) If  $f$  is strictly convex, then  $\text{argmin } f$  contains at most one point.

In other words: lsc and coercivity ensure existence of minimizers, convexity excluded non-global minimizers and, strict convexity ensures uniqueness of minimizers.

*Proof.* i) Since  $f$  is proper, we have that  $\bar{\alpha} := \inf f < \infty$ . Hence, there is a sequence  $x_n$  such that  $f(x_n) \rightarrow \bar{\alpha}$ . Since  $f$  is coercive, the sequence  $x_n$  is bounded, and hence, it has a convergent subsequence  $x_{n_k}$  with limit  $\bar{x}$ . Since  $f$  is lsc, this limit fulfills  $f(\bar{x}) \leq \liminf_k f(x_{n_k}) = \bar{\alpha}$  and this shows  $\bar{x} \in \text{argmin } f$ . Moreover  $\text{argmin } f = \text{lev}_{\bar{\alpha}} f$  and since all level sets are closed (by lsc) and bounded (by coercivity),  $\text{argmin } f$  is compact.

- ii) Let  $\bar{x}$  be a local minimizer of  $f$ , i.e. there exists  $\delta > 0$  such that  $f(\bar{x}) \leq f(x)$  for all  $x \in B_\delta(\bar{x})$ . Now let  $y \in \mathbb{R}^d$  and choose  $\lambda \in ]0, \delta / \|\bar{x} - y\|_2[$  with  $\lambda < 1$ . Then  $\lambda y + (1 - \lambda)\bar{x} \in B_\delta(\bar{x})$  and by convexity of  $f$  we get

$$f(\bar{x}) \leq f(\lambda y + (1 - \lambda)\bar{x}) \leq \lambda f(y) + (1 - \lambda)f(\bar{x}).$$

This implies  $f(\bar{x}) \leq f(y)$  and hence,  $\bar{x}$  is a global minimizer.

If  $\bar{x}, x^*$  are global minimizers, i.e.  $f(\bar{x}) = f(x^*) = \bar{\alpha}$ , then for every  $\lambda \in [0, 1]$

$$f(\lambda \bar{x} + (1 - \lambda)x^*) \leq \lambda f(\bar{x}) + (1 - \lambda)f(x^*) = \bar{\alpha},$$

i.e.  $\lambda \bar{x} + (1 - \lambda)x^*$  is also a global minimizers and hence, the set of global minimizers is convex.

- iii) If  $f$  is strictly convex, assume that  $\bar{x}$  and  $x^*$  are two different global minimizers. But then

$$f((\bar{x} + x^*)/2) < \frac{1}{2}(f(\bar{x}) + f(x^*))$$

and hence,  $(\bar{x} + x^*)/2$  would be below the global minimum which is impossible.

□

## 7 Inf-projection and inf-convolution

**Definition 7.1.** For a function  $\varphi : \mathbb{R}^d \times \mathbb{R}^m \rightarrow \bar{\mathbb{R}}$  we define the *epigraphical projection* or *inf-projection* as

$$v(u) = \inf_{x \in \mathbb{R}^d} \varphi(x, u).$$

This function is called epigraphical projection, since the epigraph of  $v$  arises as the projection of the epigraph of  $\varphi$ .

**Theorem 7.2.** If  $\varphi : \mathbb{R}^d \times \mathbb{R}^m \rightarrow \bar{\mathbb{R}}$  is convex, then its epigraphical projection  $v$  is also convex. In particular, it holds for any convex set  $C \subset \mathbb{R}^{m+1}$  that  $v(u) := \inf\{\alpha \in \mathbb{R} \mid (\alpha, u) \in C\}$  is convex.

*Proof.* For  $(x_1, u_1), (x_2, u_2) \in \text{dom } \varphi$  and  $\lambda \in [0, 1]$  we have by convexity of  $\varphi$  and the definition of  $v$  that

$$\begin{aligned} \lambda \varphi(x_1, u_1) + (1 - \lambda) \varphi(x_2, u_2) &\geq \varphi(\lambda x_1 + (1 - \lambda)x_2, \lambda u_1 + (1 - \lambda)u_2) \\ &\geq v(\lambda u_1 + (1 - \lambda)u_2). \end{aligned}$$

Taking the infimum over  $x_1$  and  $x_2$  on the left hand side gives

$$\lambda v(u_1) + (1 - \lambda)v(u_2) \geq v(\lambda u_1 + (1 - \lambda)u_2)$$

which proves the first claim. For the further claim, apply the first part to  $\varphi(\alpha, u) = \alpha + i_C(\alpha, u)$ .  $\square$

Unfortunately, epigraphical projections are not always as nice as one would want them to be:

*Example 7.3.* Let  $f : \mathbb{R} \rightarrow \mathbb{R}$  be convex and set

$$\varphi(x, u) = \begin{cases} f(x), & \exp(x) \leq u \\ \infty, & \text{else.} \end{cases}$$

This is a proper, convex function, and also lsc if  $f$  is lsc. But for  $f(x) = x$  we get that

$$v(u) = \inf_{x \in \mathbb{R}} \varphi(x, u) = \begin{cases} \infty, & u \leq 0 \\ -\infty, & u > 0 \end{cases},$$

i.e.  $v$  is neither proper nor lsc.  $\triangle$

**Definition 7.4.** For two functions  $f_1, f_2 : \mathbb{R}^d \rightarrow \bar{\mathbb{R}}$  we define the *inf-convolution* (also called *infimal convolution* or *epi-addition*) as

$$(f_1 \square f_2)(u) := \inf_{x+y=u} f_1(x) + f_2(y) = \inf_x f_1(x) + f_2(u - x)$$

If the infimum is attained whenever it is finite, we call the inf-convolution *exact*.

Since different infima commute, we get for more than two functions

$$\begin{aligned}
 f_1 \square (f_2 \square f_3)(u) &= \inf_{x+y=u} f_1(x) + (f_2 \square f_3)(y) \\
 &= \inf_{x+y=u} \left[ f_1(x) + \inf_{z+v=y} f_2(z) + f_3(v) \right] \\
 &= \inf_{\substack{x+y=u \\ z+v=y}} f_1(x) + f_2(z) + f_3(v) \\
 &= \inf_{x+z+v=u} f_1(x) + f_2(z) + f_3(v)
 \end{aligned}$$

You may compare this definition to the standard convolution of two functions  $f_1 * f_2(u) = \int f_1(x)f_2(u-x)dx$  and note that the integral (“generalized sum”) has been replaced by an infimum (“generalized minimum”) and the multiplication has been replaced by an addition.

(and we see that inf-convolutions are associative.

**Example 7.5.** 1. For two sets  $S_1$  and  $S_2$  we get for the inf-convolution of their indicator functions

$$(i_{S_1} \square i_{S_2})(u) = \inf_{x+y=u} i_{S_1}(x) + i_{S_2}(y) = i_{S_1+S_2}(u).$$

2. If we take  $f_1(x) = \|x\|$  (for some norm) and  $f_2 = i_C$  the indicator function of a convex set  $C$ , we get as their inf-convolution

$$(f_1 \square f_2)(u) = \inf_{x \in C} \|u - x\| = d(u, C),$$

i.e., the distance function for the set  $C$  (with respect to the norm  $\|\cdot\|$ )

△

**Theorem 7.6.** Let  $f_{1/2} : \mathbb{R}^d \rightarrow \bar{\mathbb{R}}$  be proper. Then it holds

$$\text{epi } f_1 + \text{epi } f_2 \subset \text{epi}(f_1 \square f_2)$$

and equality holds if and only if the inf-convolution is exact. Moreover, if  $f_1$  and  $f_2$  are convex, then so is  $f_1 \square f_2$ .

*Proof.* First, let's show the inclusion: Let  $(x, \alpha) \in \text{epi } f_1 + \text{epi } f_2$ , i.e. there exist  $(\alpha_i, x_i) \in \text{epi } f_i$  ( $i = 1, 2$ ) with  $x = x_1 + x_2$ ,  $\alpha = \alpha_1 + \alpha_2$  and  $f_1(x_1) + f_2(x_2) \leq \alpha_1 + \alpha_2 = \alpha$ . Hence

$$(f_1 \square f_2)(x) = \inf_{y_1+y_2=x} f_1(y_1) + f_2(y_2) \leq f_1(x_1) + f_2(x_2) \leq \alpha$$

and this shows that  $(x, \alpha) \in \text{epi}(f_1 \square f_2)$ .

Now let's show that there is equality precisely in the case of exactness: Let  $(x, \alpha) \in \text{epi}(f_1 \square f_2)$ , i.e.  $(f_1 \square f_2)(x) \leq \alpha$ . By exactness, there are  $x_1, x_2$  with  $x = x_1 + x_2$  and  $(f_1 \square f_2)(x) = f_1(x_1) + f_2(x_2) \leq \alpha$ . Hence

$$(x, \alpha) = (x_1, f_1(x_1)) + \underbrace{(x_2, \alpha - f_1(x_1))}_{\geq f_2(x_2)} \in \text{epi } f_1 + \text{epi } f_2.$$

For the converse implication, let  $f = f_1 \square f_2$  be finite at  $x$ , i.e.  $(x, f(x)) \in \text{epi}(f_1 \square f_2) = \text{epi } f_1 + \text{epi } f_2$ . Then, there exist  $(x_i, \alpha_i) \in \text{epi } f_i$  ( $i = 1, 2$ ) with  $(x, f(x)) = (x_1, \alpha_1) + (x_2, \alpha_2)$  and thus,  $f(x) = \alpha_1 + \alpha_2 \geq f_1(x_1) + f_2(x_2)$ . However, since we also have  $f(x) \leq f_1(x_1) + f_2(x_2)$  by the definition of the inf-convolution, we have equality and see that the infimum is attained at  $x = x_1 + x_2$ .

Finally, let  $f_1, f_2$  be convex. By Proposition 5.3, we know that  $\text{epi } f_1 + \text{epi } f_2$  is a convex set. It remains to note that

$$(f_1 \square f_2)(x) = \inf\{\alpha \in \mathbb{R} \mid (x, \alpha) \in \text{epi } f_1 + \text{epi } f_2\}$$

and Theorem 7.2 on inf-projections shows that this is convex function.  $\square$

Inf-convolutions are also not always as nice as we would like them to be:

*Example 7.7.* 1. The functions  $f_1(x) = px$  ( $p \in \mathbb{R}$ ) and  $f_2(x) = \exp(x)$  are proper, convex, and continuous. Their inf-convolution is

$$(f_1 \square f_2)(u) = \begin{cases} p(u - \log(p))p + p, & \text{if } p > 0, \\ 0, & \text{if } p = 0, \\ -\infty, & \text{if } p < 0. \end{cases}$$

Hence, for  $p < 0$ , the inf-convolution is not proper. Moreover, for  $p = 0$

$$\text{epi } f_1 = \mathbb{R} \times [0, \infty[, \quad \text{epi } f_2 = \{(x, \alpha) \mid \exp(x) \leq \alpha\}, \quad \text{epi}(f_1 \square f_2) = \mathbb{R} \times [0, \infty[$$

and hence,  $\text{epi } f_1 + \text{epi } f_2 \subset \text{epi}(f_1 \square f_2)$  with strict inclusion.

2. Consider

$$C_1 := \{(x, y) \mid y \geq \exp(x)\}, \quad C_2 := \{(x, y) \mid y \geq \exp(-x)\}$$

which are both non-empty, closed, convex sets in  $\mathbb{R}^2$  and hence, their indicator functions  $i_{C_1}$  and  $i_{C_2}$  are proper, convex and lsc. The Minkowski sum of  $C_1$  and  $C_2$  is

$$C_1 + C_2 = \mathbb{R} \times ]0, \infty[$$

which is not closed and hence, their inf-convolution  $i_{C_1} \square i_{C_2} = i_{C_1 + C_2}$  is not lsc.  $\triangle$

Interestingly, inf-convolutions are used in image processing. Here we do inf-convolution for arbitrary (bounded) functions and do not assume any convexity.

*Example 7.8* (Morphological operations). A function  $f : \mathbb{R}^d \rightarrow \mathbb{R}$  is considered as a gray scale image, i.e.  $f(x)$  denotes the gray value of the image at location  $x$ . More precisely, one would deal with function  $f : A \rightarrow [0, 1]$  where  $A$  is the image domain (e.g. a two-dimensional domain such as a rectangle) and the value  $f(x) = 0$  stands for the color black at position  $x$  while  $f(x) = 1$  stands for white.

Let  $B \subset \mathbb{R}^d$  be compact with  $B = -B$ . For a function  $f : \mathbb{R}^d \rightarrow \mathbb{R}$  we define the *erosion* of  $f$  with *structure element*  $B$  as

$$\epsilon_B f(x) = \inf_{y \in B} f(x + y) = (f \square i_B)(x)$$

and there is corresponding *dilation* defined as  $\delta_B f(x) = \sup_{y \in B} f(x + y)$ . If  $f$  represents an image, the erosion shrinks the bright areas in the image, while the dilation expands them.

The operation

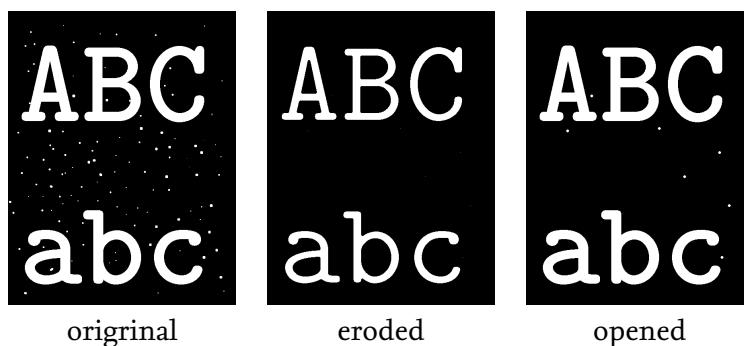
$$O_B f = \delta_{-B}(\epsilon_B f)(x)$$

is called *opening* of  $f$  and

$$C_B f = \epsilon_{-B}(\delta_B f)(x)$$

is called *closing* of  $f$  and  $\triangle$

Here is an example of an image which is degraded by white spots on dark background. The structure element is a circle (can you guess its radius?). The opening of the image eliminates the white spots but keeps the main structure (the letters) almost unchanged.



Another example is the elimination of an uneven background: To that end one calculates an image which does not contain any of the foreground structure (in this example, all dark structures should be eliminated) by a closing with a large structure element and then subtracts this background image from the original:



## 8 Proximal mappings and Moreau-Yoshida regularization

**Definition 8.1.** For a proper, convex and lsc function  $f : \mathbb{R}^d \rightarrow \bar{\mathbb{R}}$  and  $\lambda > 0$  we define the *proximal mapping*  $\text{prox}_{\lambda f} : \mathbb{R}^d \rightarrow \mathbb{R}^d$  by

$$\text{prox}_{\lambda f}(x) := \underset{y \in \mathbb{R}^d}{\operatorname{argmin}} \frac{1}{2} \|x - y\|_2^2 + \lambda f(y).$$

The *Moreau envelope* (also called *Moreau-Yoshida regularization*)  ${}^\lambda f$  is

$${}^\lambda f(x) := \min_{y \in \mathbb{R}^d} \frac{1}{2\lambda} \|x - y\|_2^2 + f(y) = \left( \frac{1}{2\lambda} \|\cdot\|_2^2 \square f \right)(y).$$

A direct consequence of the definition is that

$${}^\lambda f(x) = \frac{1}{2\lambda} \|x - \text{prox}_{\lambda f}(x)\|_2^2 + f(\text{prox}_{\lambda f}(x)).$$

Since the inf convolution of convex functions is convex, the Moreau envelope is convex, but we still need to show that both the Moreau envelope and the proximal mapping are well defined (for the first we need to show that a minimizer always exist and this will, in turn, show that the minimum for the Moreau envelope is actually assumed).

**Theorem 8.2.** For  $f : \mathbb{R}^d \rightarrow \bar{\mathbb{R}}$  proper, convex and lsc and  $\lambda > 0$  it holds that:

1. For every  $x$  there exists a unique minimizer  $\hat{x} = \text{prox}_{\lambda f}(x)$  of  $\frac{1}{2} \|x - y\|_2^2 + \lambda f(y)$ .
2. This  $\hat{x}$  is characterized by the variational inequality
$$\forall y \in \mathbb{R}^d : \langle x - \hat{x}, y - \hat{x} \rangle + \lambda(f(\hat{x}) - f(y)) \leq 0.$$
3. Some  $x^*$  is a minimizer of  $f$  if and only if  $x^*$  is a fixed point of  $\text{prox}_{\lambda f}$  for any  $\lambda > 0$ , i.e.  $x^* = \text{prox}_{\lambda f}(x^*)$ .
4. The Moreau envelope  ${}^\lambda f$  is differentiable with gradient

$$\nabla({}^\lambda f)(x) = \frac{1}{\lambda}(x - \text{prox}_{\lambda f}(x)).$$

5. It holds that

$$\underset{x}{\operatorname{argmin}} f(x) = \underset{x}{\operatorname{argmin}} {}^\lambda f(x).$$

*Proof.* 1. The objective in the definition of the proximal mapping is strictly convex and hence, has at most one minimizer. Let us now show existence of a minimizer: By Proposition 5.4 there exists an affine lower bound for  $f$ , i.e. we have  $f(y) \geq \langle p, y \rangle + \alpha$  for some  $p$  and  $\alpha$ . Hence, we have

$$\begin{aligned} g_x(y) &:= \frac{1}{2} \|y - x\|_2^2 + \lambda f(y) \geq \frac{1}{2} \|y - x\|_2^2 + \lambda \langle p, y \rangle + \lambda \alpha \\ &= \frac{1}{2} \|y - x + \lambda p\|_2^2 + \lambda \langle p, x \rangle - \frac{\lambda^2}{2} \|p\|^2 + \lambda \alpha. \end{aligned}$$

Hence, the function on the left is coercive and since it also lsc, we see by Theorem 6.5, that minimizers exist.

Note that we could have defined the proximal mapping differently by taking the argmin of  $\frac{1}{2\lambda} \|x - y\|_2^2 + f(y)$ , i.e. the proximal mapping maps  $x$  to the point where the minimum in the Moreau envelope is assumed.

We could also have observed that  $g_x$  is strongly convex and use an exercise from Sheet 4 to deduce coercivity.

2. If  $\hat{x} = \text{prox}_{\lambda f}(x)$ , then for every  $\mu \in [0, 1]$  and  $y$

$$g_x(\hat{x}) \leq g_x(\hat{x} + \mu(y - \hat{x}))$$

which we rearrange to

$$\begin{aligned} 0 &\geq \lambda(f(\hat{x}) - f(\hat{x} + \mu(y - \hat{x}))) + \frac{1}{2}(\|\hat{x} - x\|_2^2 - \|\hat{x} - x + \mu(y - \hat{x})\|_2^2) \\ &\geq \lambda(f(\hat{x}) - f(\hat{x} + \mu(y - \hat{x}))) + \frac{1}{2}(-\mu^2\|y - \hat{x}\|_2^2 - 2\mu\langle \hat{x} - x, y - \hat{x} \rangle) \end{aligned}$$

By convexity of  $f$  we get

$$0 \geq \lambda\mu(f(\hat{x}) - f(y)) + \mu\langle \hat{x} - x, \hat{x} - y \rangle - \frac{\mu^2}{2}\|\hat{x} - y\|_2^2$$

Dividing by  $\mu$  and letting  $\mu \rightarrow 0$  shows the claim.

Conversely, let the variational inequality be fulfilled. This means that for all  $y \in \mathbb{R}^d$

$$\begin{aligned} \lambda f(\hat{x}) + \frac{1}{2}\|\hat{x} - x\|_2^2 &\leq \lambda f(y) - \langle \hat{x} - x, \hat{x} - y \rangle + \frac{1}{2}\|\hat{x} - x\|_2^2 \\ &\leq \lambda f(y) + \frac{1}{2}\|\hat{x} - y\|_2^2 - \langle \hat{x} - x, \hat{x} - y \rangle + \frac{1}{2}\|\hat{x} - x\|_2^2 \\ &= \lambda f(y) + \frac{1}{2}\|y - x\|_2^2 \end{aligned}$$

and this shows  $\hat{x} = \text{prox}_{\lambda f}(x)$ .

3. If  $\hat{x}$  is a minimizer of  $f$ , then  $f(\hat{x}) \leq f(x)$  for all  $x$  and hence  $\lambda f(\hat{x}) + \frac{1}{2}\|\hat{x} - \hat{x}\|_2^2 \leq \lambda f(x) + \frac{1}{2}\|x - \hat{x}\|_2^2$  which shows that  $\hat{x} = \text{prox}_{\lambda f}(\hat{x})$ .

Conversely, if  $\hat{x} = \text{prox}_{\lambda f}(\hat{x})$ , we get, by 2., that for all  $y$

$$\langle \hat{x} - \hat{x}, y - \hat{x} \rangle + \lambda(f(\hat{x}) - f(y)) \leq 0$$

and this shows  $f(\hat{x}) \leq f(y)$  for all  $y$ .

4. For  $x_0 \in \mathbb{R}^d$  define  $\hat{x}_0 := \text{prox}_{\lambda f}(x_0)$  and  $z := \frac{x_0 - \hat{x}_0}{\lambda}$ . To show that  ${}^\lambda f$  is differentiable at  $x_0$  with  $({}^\lambda f)'(x_0) = z$  we need to show that

$$r(u) := {}^\lambda f(x_0 + u) - {}^\lambda f(x_0) - \langle z, u \rangle = o(\|u\|_2).$$

By the definition of the prox and the Moreau envelope, we have

$$\begin{aligned} {}^\lambda f(x_0) &= f(\hat{x}_0) + \frac{1}{2\lambda}\|\hat{x}_0 - x_0\|_2^2 \\ {}^\lambda f(x_0 + u) &= \min_x \left[ f(x) + \frac{1}{2\lambda}\|x - x_0 - u\|_2^2 \right] \\ &\leq f(\hat{x}_0) + \frac{1}{2\lambda}\|\hat{x}_0 - x_0 - u\|_2^2. \end{aligned}$$

It follows (plugging in  $z$ )

$$r(u) \leq \frac{1}{2\lambda}\|\hat{x}_0 - x_0 - u\|_2^2 - \frac{1}{2\lambda}\|\hat{x}_0 - x_0\|_2^2 - \frac{1}{\lambda}\langle x_0 - \hat{x}_0, u \rangle = \frac{1}{2\lambda}\|u\|_2^2.$$



Using the convexity of  $r$ , we get

$$0 = r(0) = r\left(\frac{1}{2}u + \frac{1}{2}(-u)\right) \leq \frac{1}{2}(r(u) + r(-u))$$

and this shows

$$r(u) \geq -r(-u) \geq -\frac{1}{2\lambda}\|u\|_2^2$$

and in total we get  $|r(u)| \leq \frac{1}{2\lambda}\|u\|_2^2$  which proves the claim.

5. We know by 3. that  $\hat{x}$  is a minimizer of  $f$  if and only if  $\hat{x} = \text{prox}_{\lambda f}(\hat{x})$  and by 4. this is equivalent to  $\nabla(\lambda f)(\hat{x}) = 0$ , i.e. exactly when  $\hat{x}$  is a minimizer of  $\lambda f$ .

□

**Example 8.3** (Prox of indicators are projections). Let  $f(y) = i_C(y)$  for some non-empty, convex and closed  $C$ . Then (note that  $\lambda i_C = i_C$ )

$$\text{prox}_{i_C}(x) = \underset{y \in \mathbb{R}^d}{\text{argmin}} \frac{1}{2}\|x - y\|_2^2 + \lambda i_C(y) = \underset{y \in C}{\text{argmin}} \|x - y\| = P_C(x),$$

i.e.  $\text{prox}_{\lambda f}(x)$  is the orthogonal projection of  $x$  onto  $C$  (independently of  $\lambda > 0$ ).

The Moreau envelope is

$$\lambda i_C(x) = \frac{1}{2\lambda}\|x - P_C(x)\|_2^2 = \frac{1}{2\lambda}d(x, C)^2.$$

△

**Example 8.4** (Prox of  $|\cdot|$  is soft shrinkage). Now consider  $f(x) = |x|$ . To calculate  $\hat{x} := \underset{y}{\text{argmin}} \frac{1}{2}(x - y)^2 + \lambda|y|$  we first observe that if  $x \geq 0$ , then  $\hat{x} \geq 0$  and also  $\text{prox}_{\lambda f}(-x) = -\text{prox}_{\lambda f}(x)$ . Hence, we only consider  $x > 0$  and have

$$\hat{x} = \underset{y \geq 0}{\text{argmin}} \frac{1}{2}(x - y)^2 + \lambda y.$$

We note that the unconstrained minimizer is  $x - \lambda$ , and if  $0 \leq x \leq \lambda$ , then the minimizer is 0. In total, we get that

$$\begin{aligned} \text{prox}_{\lambda|\cdot|}(x) &= \begin{cases} x - \lambda, & \text{if } x > \lambda, \\ 0, & \text{if } |x| \leq \lambda, \\ x + \lambda, & \text{if } x < -\lambda, \end{cases} \\ &= \max(|x| - \lambda, 0) \text{sign}(x) =: S_\lambda(x). \end{aligned}$$

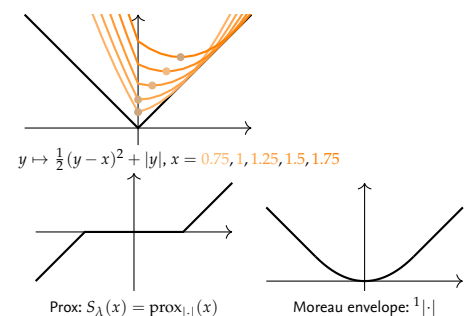
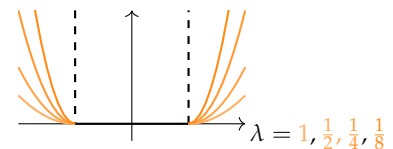
This function is known as the *soft thresholding* or *soft shrinkage* function.

To calculate the Moreau envelope of  $|\cdot|$  we just plug in and get

$$\begin{aligned} \lambda f(x) &= \frac{1}{2\lambda}(x - S_\lambda(x))^2 + |S_\lambda(x)| \\ &= \begin{cases} |x| - \frac{\lambda}{2}, & \text{if } |x| > \lambda \\ \frac{1}{2\lambda}x^2, & \text{if } |x| \leq \lambda. \end{cases} \end{aligned}$$

This function is called *Huber function*.

△



**Lemma 8.5** (Calculus for proximal mappings). *For proper, convex and lsc functions  $f$  and  $\lambda > 0$  it holds that:*

- i) If  $g(x) = f(x) + \alpha$  for  $\alpha \in \mathbb{R}$ , then  $\text{prox}_{\lambda g} = \text{prox}_{\lambda f}$ .
- ii) If  $g(x) = \tau f(\mu x)$  for  $\tau > 0, \mu \in \mathbb{R}$ , then  $\text{prox}_{\lambda g}(x) = \frac{1}{\mu} \text{prox}_{\lambda \mu^2 \tau f}(\mu x)$ .
- iii) If  $g(x) = f(x + x_0) + \langle p, x \rangle$ , for  $x_0, p \in \mathbb{R}^d$ , then  $\text{prox}_{\lambda g}(x) = \text{prox}_{\lambda f}(x + x_0 + \lambda p) - x_0$ .
- iv) If  $g(x) = f(Qx)$  for orthonormal  $Q \in \mathbb{R}^{d \times d}$ , then  $\text{prox}_{\lambda g}(x) = Q^T \text{prox}_{\lambda f}(Qx)$ .
- v) If  $h(x, y) = f(x) + g(y)$  for proper, convex and lsc  $g$ , then  $\text{prox}_{\lambda h}(x, y) = (\text{prox}_{\lambda f(x), \text{prox}_{\lambda g}(y)})$ .

*Proof.* The first three items are straightforward implications from the definition by appropriate substitutions. For item iv) we write

$$\text{prox}_{\lambda g}(x) = \underset{y}{\operatorname{argmin}} \frac{1}{2} \|x - y\|_2^2 + \lambda f(Qy)$$

and use that  $Q^T$  is onto by assumption. Hence, we can substitute  $y = Q^T z$ , minimize over  $z$ , but keep in mind that we are interested in the *minimizer*. Hence, we have, using  $QQ^T = I$

$$\begin{aligned} \text{prox}_{\lambda g}(x) &= Q^T \underset{z}{\operatorname{argmin}} \frac{1}{2} \|x - Q^T z\|_2^2 + \lambda f(z) \\ &= Q^T \underset{z}{\operatorname{argmin}} \frac{1}{2} \|Qx - z\|_2^2 + \lambda f(z) \end{aligned}$$

where we used that  $\|Qu\|_2 = \|u\|_2$  since  $Q$  is orthonormal. For item v) just observe that the minimization can be carried out independently over  $x$  and  $y$ .

□

## 9 Subgradients

If you've found it surprising that convex functions automatically have some continuity property, you may find it even more surprising, that they also have some inbuilt differentiability.

We recall the notion of directional derivative: For  $f : \mathbb{R}^d \rightarrow \mathbb{R}$ ,  $x_0, v \in \mathbb{R}^d$  we denote by

$$Df(x_0, v) := \lim_{h \searrow 0} \frac{f(x_0 + hv) - f(x_0)}{h}$$

the *one-sided directional derivative* (whenever the limit exists).

**Definition 9.1.** A function  $f : \mathbb{R}^d \rightarrow \bar{\mathbb{R}}$  is *positively homogeneous* (of degree 1) if

$$\lambda > 0 \implies f(\lambda x) = \lambda f(x).$$

Norms and semi-norms are obvious examples of positive homogeneous functions as well as linear functions.

**Theorem 9.2.** Let  $f : \mathbb{R}^d \rightarrow \bar{\mathbb{R}}$  be convex,  $x_0 \in \text{dom}(f)$ ,  $v \in \mathbb{R}^d$ . Then, the limit in the definition of the directional derivative exists in  $\bar{\mathbb{R}}$  and it holds

$$Df(x_0, v) = \inf_{h > 0} \frac{f(x_0 + hv) - f(x_0)}{h}.$$

Moreover, the function  $Df(x_0, \cdot)$  is convex and positively homogeneous and for any  $v$  it holds that

$$Df(x_0, v) \leq f(x_0 + v) - f(x_0)$$

and if  $f(x_0 - v) \neq -\infty$  it also holds that

$$f(x_0) - f(x_0 - v) \leq Df(x_0, v).$$

In particular,  $Df(x_0, v)$  is finite for all  $v$  if  $x_0 \in \text{int dom}(f)$ .

*Proof.* The proof relies on the following monotonicity property of the difference quotient: For  $0 < h_1 < h_2$  it holds that

$$\frac{f(x_0 + h_1 v) - f(x_0)}{h_1} \leq \frac{f(x_0 + h_2 v) - f(x_0)}{h_2}.$$

To see this first note that if  $x_0 + h_2 v \notin \text{dom } f$ , the inequality is always fulfilled. Hence, we assume that  $x_0 + h_2 v \in \text{dom } f$ . By convexity of  $f$  we get

$$\begin{aligned} f(x_0 + h_1 v) - f(x_0) &= f\left(\frac{h_1}{h_2}(x_0 + h_2 v) + \left(1 - \frac{h_1}{h_2}\right)x_0\right) - f(x_0) \\ &\leq \frac{h_1}{h_2}f(x_0 + h_2 v) + \left(1 - \frac{h_1}{h_2}\right)f(x_0) - f(x_0) \\ &= \frac{h_1}{h_2}(f(x_0 + h_2 v) - f(x_0)), \end{aligned}$$

and this shows the desired monotonicity.

Recall that in the case of differentiable  $f$ , it holds that  $Df(x_0, v) = \langle \nabla f(x_0), v \rangle$ , but that directional differentiability in all directions does not imply differentiability.

Positive homogeneity of degree  $p$  would mean  $f(\lambda x) = \lambda^p f(x)$ , but we will not need this notion.

The monotonicity implies that the limit in the definition of the directional derivative is actually an infimum and, moreover,  $Df(x_0, v) \leq f(x_0 + v) - f(x_0)$  follows by taking  $h = 1$ .

To show the lower inequality, let  $v$  be such that  $f(x_0 - v) > -\infty$ . If  $f(x_0 - v) = \infty$  or  $Df(x_0, v) = \infty$ , there is nothing to prove. Hence, assume that both quantities are finite. Since  $\infty > Df(x_0, v) = \lim_{h \rightarrow 0} (f(x_0 + hv) - f(x_0))/h$ , there is an  $\bar{h} > 0$  such that  $\frac{f(x_0 + hv) - f(x_0)}{h}$  is finite for all  $h \in ]0, \bar{h}]$ . Hence,  $x_0 - v, x_0 + hv \in \text{dom } f$  for these  $h$  and we get, again using convexity,

$$\begin{aligned} f(x_0) &= f\left(\frac{1}{1+h}(x_0 + hv) + \frac{h}{1+h}(x_0 - v)\right) \\ &\leq \frac{1}{1+h}f(x_0 + hv) + \frac{h}{1+h}f(x_0 - v). \end{aligned}$$

We rearrange to

$$f(x_0) - f(x_0 - v) \leq \frac{f(x_0 + hv) - f(x_0)}{h}$$

and the claim follows for  $h \searrow 0$ .

Now we show that  $Df(x_0, \cdot)$  is convex: For  $v_1, v_2 \in \text{dom } Df(x_0, \cdot)$  we have (for small enough  $h$  and  $\lambda \in [0, 1]$ )

$$\begin{aligned} f(x_0 + h(\lambda v_1 + (1 - \lambda)v_2)) - f(x_0) &= f(\lambda(x_0 + hv_1) + (1 - \lambda)(x_0 + hv_2)) - \lambda f(x_0) - (1 - \lambda)f(x_0) \\ &\leq \lambda(f(x_0 + hv_1) - f(x_0)) + (1 - \lambda)(f(x_0 + hv_2) - f(x_0)) \end{aligned}$$

Dividing by  $h > 0$  and  $h \searrow 0$  shows the desired convexity.

To show that  $Df(x_0, \cdot)$  is positively homogeneous, we observe for  $\lambda > 0$

$$\begin{aligned} Df(x_0, \lambda v) &= \lim_{h \searrow 0} \frac{f(x_0 + \lambda hv) - f(x_0)}{h} \\ &= \lim_{\bar{h} \searrow 0} \lambda \frac{f(x_0 + \bar{h}v) - f(x_0)}{\bar{h}} \quad (\bar{h} = \lambda h) \\ &= \lambda Df(x_0, v) \end{aligned}$$

as desired.

Finally, we show that  $Df(x_0, v)$  is finite for  $x_0 \in \text{int dom } f$  and all  $v$ : For every  $v$  there is  $\lambda > 0$  such that  $x_0 \pm \lambda v \in \text{int dom } f$  and thus

$$\frac{1}{\lambda}(f(x_0) - f(x_0 - \lambda v)) \leq \underbrace{\frac{1}{\lambda} Df(x_0, \lambda v)}_{= Df(x_0, v)} \leq \frac{1}{\lambda}(f(x_0 + \lambda v) - f(x_0))$$

and the claim follows since the leftmost and rightmost expressions are finite.  $\square$

Now recall that by Theorem 5.1 we know that a differentiable and convex  $f$  fulfills that for every  $x_0$  it holds that

$$\forall x : f(x) \geq f(x_0) + \langle \nabla f(x_0), x - x_0 \rangle.$$

**Definition 9.3.** Let  $f : \mathbb{R}^d \rightarrow \bar{\mathbb{R}}$  be convex and  $x_0 \in \text{dom } f$ . Then  $p \in \mathbb{R}^d$  is a *subgradient* of  $f$  at  $x_0$  if

$$\forall x \in \mathbb{R}^d : f(x) \geq f(x_0) + \langle p, x - x_0 \rangle.$$

The set of all subgradients of  $f$  at  $x_0$  is denoted by  $\partial f(x_0)$  and called *subdifferential* of  $f$  at  $x_0$ .

If  $x_0 \notin \text{dom } f$  we set  $\partial f(x_0) = \emptyset$ . Furthermore we denote

$$\text{dom } \partial f = \{x \mid \partial f(x) \neq \emptyset\}$$

and  $f$  is called *subdifferentiable* at  $x_0$  if  $x_0 \in \text{dom } \partial f$ .

A direct consequence of the definition is that  $\partial f(x_0)$  is always closed and convex.

If  $p_n \in \partial f(x)$  with  $p_n \rightarrow p$ , then  $f(y) \geq f(x) + \langle p_n, y - x \rangle$  for all  $y$  and passing to the limit  $p_n \rightarrow p$  shows that  $p \in \partial f(x)$ . If  $p, q \in \partial f(x)$ , then  $f(y) = \lambda f(y) + (1 - \lambda)f(y) \geq \lambda f(x) + \lambda \langle p, y - x \rangle + (1 - \lambda)f(x) + (1 - \lambda) \langle q, y - x \rangle = f(x) + \langle \lambda p + (1 - \lambda)q, y - x \rangle$ , i.e.  $\lambda p + (1 - \lambda)q \in \partial f(x)$ .

Graphically, some  $p$  is a subgradient of  $f$  at  $x_0$  if the function  $x \mapsto f(x_0) + \langle p, x - x_0 \rangle$  is a supporting affine lower bound which is exact at  $x_0$ .

The nice thing about subgradient is, that they even exist where a convex function has a kink, and in this case there is more than one of them:

*Example 9.4* (Subgradients of the absolute value). The function  $f(x) = |x|$  on  $\mathbb{R}$  fulfills

$$\partial f(x) = \begin{cases} \{1\}, & \text{if } x > 0, \\ [-1, 1], & \text{if } x = 0, \\ \{-1\}, & \text{if } x < 0, \end{cases}$$

which can be seen by direct inspection.  $\triangle$

**Proposition 9.5** (The subdifferential and directional derivatives). Let  $f : \mathbb{R}^d \rightarrow \bar{\mathbb{R}}$  be convex and  $f(x_0) \in \mathbb{R}$ . Then it holds that

$$\partial f(x_0) = \{p \in \mathbb{R}^d \mid \forall v \in \mathbb{R}^d : \langle p, v \rangle \leq Df(x_0, v)\}. \quad (*)$$

*Proof.* Denote the set on the right hand side of (\*) by  $M$ .

$\partial f(x_0) \subset M$ : Let  $p \in \partial f(x_0)$  and set  $x = x_0 + hv$  for  $h > 0$ . Then, by definition of the subdifferential:

$$f(x_0 + hv) - f(x_0) \geq \langle p, hv \rangle$$

and division by  $h$  and  $h \searrow 0$  shows  $p \in M$ .

$\partial f(x_0) \supset M$ : Now let  $p \in M$ . Theorem 9.2 shows that for all  $v \in \mathbb{R}^d$ .

$$\langle p, v \rangle \leq Df(x_0, v) \leq f(x_0 + v) - f(x_0)$$

and the claim follows if we set  $x = x_0 + v$

□

Subgradients have some properties similar to gradients:

**Proposition 9.6.** Let  $f : \mathbb{R}^d \rightarrow \bar{\mathbb{R}}$  be convex. Then it holds:

- i)  $\partial\varphi(x) = \partial f(x + x_0)$  for  $\varphi(x) := f(x + x_0)$ .
- ii)  $\partial\varphi(x) = \lambda\partial f(\lambda x)$  for  $\varphi(x) := f(\lambda x)$ .
- iii)  $\partial\varphi(x) = \lambda\partial f(x)$  for  $\varphi(x) := \lambda f(x)$ .

These are straightforward consequences of the definition and one should do the proofs as exercises.

**Proposition 9.7.** A convex function  $f : \mathbb{R}^d \rightarrow \bar{\mathbb{R}}$  is differentiable at  $x_0$  if and only if  $f$  is continuous at  $x_0$  and  $\partial f(x_0)$  only has one element and in this case we have  $\partial f(x_0) = \{\nabla f(x_0)\}$ .

*Proof.* Let  $f$  be differentiable at  $x_0$ . By the previous lemma we have that  $p \in \partial f(x_0)$  fulfills  $\langle p, v \rangle \leq \langle \nabla f(x_0), v \rangle$  for all  $v$  which implies  $p = \nabla f(x_0)$ .

To show the converse, we assume without loss of generality that  $x_0 = 0$  and  $\partial f(0) = \{w\}$ . We define the convex function  $F_v(t) = f(tv)$ . Similarly to Proposition 9.6 ii) we see that  $\partial F_v(0) = \{\langle w, v \rangle\}$  which implies for  $t > 0$

$$0 \leq \frac{F_v(t) - F_v(0)}{t} - \langle w, v \rangle.$$

On the other hand, for  $\epsilon > 0$  there exists  $t_\epsilon$  such that  $F_v(t_\epsilon) < F_v(0) + t_\epsilon \langle w, v \rangle + t_\epsilon \epsilon$  (note that  $\langle w, v \rangle + \epsilon \notin \partial F_v(0)$ ). Convexity of  $F_v$  shows that for every  $t \in [0, t_\epsilon]$  it holds that

$$\begin{aligned} F_v(t) &\leq \frac{t}{t_\epsilon} F_v(t_\epsilon) + \frac{t_\epsilon - t}{t_\epsilon} F_v(0) \\ &\leq F_v(0) + t \langle w, v \rangle + t \epsilon, \end{aligned}$$

and hence,

$$\frac{F_v(t) - F_v(0)}{t} - \langle w, v \rangle \leq \epsilon.$$

Since  $\epsilon > 0$  was arbitrary, it follows that

$$\lim_{t \searrow 0} \frac{F_v(t) - F_v(0)}{t} = \langle w, v \rangle$$

which proves differentiability of  $f$  and  $\nabla f(0) = w$ . □

**Example 9.8.** 1. We begin with the very simple example  $f(x) = |x|$  on the real line. In one dimension, we can characterize the subdifferential completely by left- and right-derivatives, since these are the directional derivatives  $Df(x, -1)$  and  $Df(x, 1)$ , respectively. For the absolute value we get

$$\partial f(x) = \begin{cases} \{-1\}, & \text{if } x < 0 \\ [-1, 1], & \text{if } x = 0 \\ \{1\}, & \text{if } x > 0. \end{cases}$$

2. Now consider the function

$$f(x) = \begin{cases} \infty, & \text{if } x < 0 \text{ or } x > 1 \\ -\sqrt{x}, & \text{if } 0 \leq x \leq 1. \end{cases}$$

We get

$$\partial f(x) = \begin{cases} \emptyset, & \text{if } x \leq 0, \\ \{-\frac{1}{2\sqrt{x}}\}, & \text{if } 0 < x < 1, \\ [-\frac{1}{2}, \infty[, & \text{if } x = 1, \\ \emptyset, & \text{if } x > 1. \end{cases}$$

Note that the subdifferential  $\partial f(x_0)$  can be empty or unbounded if  $x_0 \notin \text{int dom } f$ . The next theorem shows that none of this can happen in the interior of the domain.

△

## 10 Subdifferential calculus

Proper and convex functions always have subgradients in the interior of their domain.

**Theorem 10.1.** *If  $f : \mathbb{R}^d \rightarrow \bar{\mathbb{R}}$  be proper and convex. Then it holds that  $\partial f(x_0)$  is non-empty and bounded if  $x_0 \in \text{int dom } f$ .*

*Proof.* If  $x_0 \in \text{int dom } f$ , then  $f$  is locally Lipschitz-continuous at  $x_0$  (Lemma 6.1), hence, there is  $\delta > 0$  such that  $|f(x) - f(x_0)| < 1$  for  $\|x - x_0\| < \delta$ . In particular, for some  $p \in \partial f(x_0)$  and all  $x$  with  $\|x\|_2 < 1$  we have

$$1 > f(x_0 + \delta x) - f(x_0) \geq \langle p, \delta x \rangle.$$

Hence,  $\langle p, x \rangle \leq 1/\delta$ , and taking the supremum over all  $x$  with  $\|x\|_2 < 1$  shows  $\|p\|_2 \leq 1/\delta$ .

To show that  $\partial f(x_0)$  is non-empty, note that  $\text{int}(\text{epi } f)$  is not empty (since it contains an open set of the form  $B_\delta(x_0) \times ]f(x_0), \infty[$ ). Moreover  $(x_0, f(x_0))$  is not in  $\text{int}(\text{epi } f)$  and hence we can, by Proposition 3.7, find  $(p_0, t) \in \mathbb{R}^d \times \mathbb{R}$  for fulfills  $(p_0, t) \neq (0, 0)$  and

$$\langle p_0, x \rangle + ts \leq \lambda \quad \text{if } x \in \text{dom } f, f(x) \leq s \quad \text{and} \quad \langle p_0, x_0 \rangle + tf(x_0) \geq \lambda.$$

With  $x = x_0$  and  $s = f(x_0)$  we see that  $\lambda = \langle p_0, x_0 \rangle + tf(x_0)$ . Moreover, we see that  $t < 0$ , since  $t > 0$  would lead to a contradiction by sending  $t \rightarrow \infty$  and  $t = 0$  would imply  $\langle p_0, x - x_0 \rangle \leq 0$  for all  $x \in B_\delta(x_0)$  and this would imply  $p_0 = 0$  which would contradict  $(p_0, t) \neq 0$ . With  $p = -p_0/t$  we get

$$f(x_0) + \langle p, x - x_0 \rangle \leq f(x_0)$$

for all  $x \in \text{dom } f$  and this shows  $p \in \partial f(x_0)$ .  $\square$

Derivatives can be used to find local minima of functions. For convex functions, this is accomplished by subgradients. However, due to convexity, the first order condition is necessary and sufficient.

**Theorem 10.2** (Fermat's rule). *Let  $f : \mathbb{R}^d \rightarrow \bar{\mathbb{R}}$  be proper and convex. Then it holds that  $\hat{x}$  is a global minimizer of  $f$  if and only if  $0 \in \partial f(\hat{x})$ .*

*Proof.* Let  $\hat{x}$  be a global minimizer. Hence, we have for all  $x$

$$\begin{aligned} f(x) &\geq f(\hat{x}) \\ &= f(\hat{x}) + \langle 0, x - \hat{x} \rangle \end{aligned}$$

and this shows  $0 \in \partial f(\hat{x})$ . Conversely, if  $0 \in \partial f(\hat{x})$ , then the above subgradient inequality holds for all  $x$  and hence,  $\hat{x}$  is a global minimizer.  $\square$

There is sharper result, namely that the subdifferential  $\partial f(x_0)$  is non-empty and bounded if and only if  $x_0 \in \text{ri dom } f$ , but we will not prove this here.



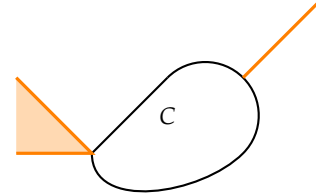
We close this section with another important example:

**Example 10.3.** Let  $C$  be convex and non-empty. Then the subdifferential of the indicator function of  $C$  is given by

$$\partial i_C(x_0) = \begin{cases} \{p \in \mathbb{R}^d \mid \forall x \in C : \langle p, x - x_0 \rangle \leq 0\}, & \text{if } x_0 \in C \\ \emptyset, & \text{if } x_0 \notin C \end{cases}$$

First, we observe, that for  $x_0 \in \text{int } C$ , it holds that  $\partial i_C(x_0) = \{0\}$  (as the function is locally constant there). If  $x \in \partial C$ , then there is geometric meaning of the subgradients: A vector  $p$  is a subgradient, if the angle between  $p$  and the line from  $x_0$  to any point  $x \in C$  is larger than  $90^\circ$ . We have seen this property already and indeed it holds that the subgradient of the indicator is in fact the normal cone, i.e.

$$\partial i_C(x) = N_C(x).$$



△

However, the situation is a bit more subtle for the sum-rule and the concatenation with general linear operators:

**Theorem 10.4.** i) Let  $f, g : \mathbb{R}^d \rightarrow \bar{\mathbb{R}}$  be proper and convex. Then it holds for every  $x$  that

$$\partial f(x) + \partial g(x) \subset \partial(f + g)(x).$$

Equality holds if there exists some  $\bar{x}$  such that  $\bar{x} \in \text{dom}(f) \cap \text{dom}(g)$  and  $f$  is continuous at  $\bar{x}$ .

ii) Let  $f : \mathbb{R}^d \rightarrow \bar{\mathbb{R}}$  be proper and convex and  $A \in \mathbb{R}^{d \times n}$ . Then  $\varphi(x) := f(Ax)$  satisfies

$$A^T \partial f(Ax) \subset \partial \varphi(x)$$

for all  $x$  and equality holds if there exists  $\bar{x}$  such that  $f$  is continuous and finite at  $A\bar{x}$ .

**Proof.** i) We start with the inclusion: Let  $p \in \partial f(x) + \partial g(x)$ , i.e. we have  $p = q + r$  with  $q \in \partial f(x)$ ,  $r \in \partial g(x)$ . Hence, we have for all  $y$  that

$$\begin{aligned} f(y) &\geq f(x) + \langle q, y - x \rangle \\ g(y) &\geq g(x) + \langle r, y - x \rangle \end{aligned}$$

and adding these inequalities shows that  $p = q + r \in \partial(f + g)(x)$ .

Now assume that  $f$  is continuous at some  $\bar{x}$  with  $\bar{x} \in \text{dom}(f) \cap \text{dom}(g)$  and it remains to show that  $\partial(f + g)(x) \subset \partial f(x) + \partial g(x)$ . If  $f(x) = \infty$  or  $g(x) = \infty$ , then the inclusion is clear, since then  $\partial(f + g)(x) = \emptyset$ .

Note that the point  $\bar{x}$  is totally unrelated to the  $x$  in the assertion and that only one of the two functions need to be continuous at  $\bar{x}$ .

Hence, let  $x \in \text{dom}(f) \cap \text{dom}(g)$ . We have to prove that every  $p \in \partial(f+g)(x)$  can be decomposed into  $p = q + r$  with  $q \in \partial f(x)$  and  $r \in \partial g(x)$ . The subgradient inequality for  $p \in \partial(f+g)(x)$  is

$$f(y) + g(y) \geq f(x) + g(x) + \langle p, y - x \rangle$$

for all  $y$ . We rewrite this as

$$F(y) := f(y) - f(x) - \langle p, y - x \rangle \geq g(x) - g(y) =: G(y) \quad (*)$$

and consider the sets

$$C_1 := \{(y, \alpha) \in \mathbb{R}^d \times \mathbb{R} \mid \alpha \geq F(y)\} = \{(y, \alpha) \in \text{dom}(f) \times \mathbb{R} \mid \alpha \geq F(y)\}$$

$$C_2 := \{(y, \alpha) \in \mathbb{R}^d \times \mathbb{R} \mid G(y) \geq \alpha\} = \{(y, \alpha) \in \text{dom}(g) \times \mathbb{R} \mid G(y) \geq \alpha\}.$$

Since  $F(x) = G(x) = 0$ , both sets are non-empty and since  $f$  and  $g$  are convex (hence  $F$  is convex and  $G$  is concave), both sets are convex. By (\*), the sets  $C_1$  and  $C_2$  only share boundary points. Hence, by the proper separation theorem (Theorem 3.9), there exists a nonzero  $(q, a)$  such that

$$\left\langle \begin{bmatrix} q \\ a \end{bmatrix}, \begin{bmatrix} y_2 \\ \alpha_2 \end{bmatrix} \right\rangle \leq \left\langle \begin{bmatrix} q \\ a \end{bmatrix}, \begin{bmatrix} y_1 \\ \alpha_1 \end{bmatrix} \right\rangle$$

for  $(y_i, \alpha_i) \in C_i, i = 1, 2$ .

We aim to show  $a > 0$ : Since  $(x, \alpha_1) \in C_1$  for  $\alpha_1 \geq 0$  and  $(x, \alpha_2) \in C_2$  for  $\alpha_2 \leq 0$ , we have that  $a\alpha_2 \leq a\alpha_1$ . Hence,  $a \geq 0$ .

Now we show that  $a > 0$ , i.e. we only need to show that  $a \neq 0$ . If  $a = 0$  would hold, then we would have

$$\langle q, y_2 \rangle \leq \langle q, y_1 \rangle, \quad \text{for } y_1 \in \text{dom } F = \text{dom } f, y_2 \in \text{dom}(-G) = \text{dom } g$$

$$\langle q, \bar{x} \rangle \leq \langle q, y_1 \rangle \quad \text{for } y_1 \in \text{dom } F = \text{dom } f.$$

In particular, the continuity of  $f$  would imply that for  $y_1 = \bar{x} \pm \Delta x \in \text{dom}(f)$  (which holds for all  $\Delta x$  small enough), that  $0 \leq \langle q, \pm \Delta x \rangle$  for all these  $\Delta x$  and this would imply  $q = 0$  which contradicts  $(q, a) \neq 0$ .

Thus  $a > 0$  and we can assume  $a = 1$  without loss of generality. We have for  $(y, F(y)) \in C_1$  and  $(y, G(y)) \in C_2$  that whenever  $y$  is in  $\text{dom}(f)$  or  $\text{dom}(g)$ , respectively, that

$$\langle q, y \rangle + F(y) \geq b \quad \langle q, y \rangle + G(y) \leq b.$$

With  $y = x \in \text{dom}(f) \cap \text{dom}(g)$ , we get the equality  $\langle q, x \rangle = b$  and we get

$$\forall y \in \text{dom}(f) : f(y) - f(x) - \langle p, y - x \rangle + \langle q, y \rangle \geq \langle q, x \rangle$$

which means

$$\forall y \in \text{dom}(f) : f(y) \geq f(x) + \langle p - q, y - x \rangle.$$

We conclude  $p - q \in \partial f(x)$  and similarly, we get  $q \in \partial g(x)$  which proves the claim.

ii) For the inclusion let  $q = A^T p$  with  $p \in \partial f(Ax)$  we have for all  $y$  that

$$\begin{aligned}\varphi(x) + \langle q, y - x \rangle &= f(Ax) + \langle A^T p, y - x \rangle \\ &= f(Ax) + \langle p, Ay - Ax \rangle \leq f(Ay) = \varphi(y)\end{aligned}$$

and this shows that  $q \in \partial\varphi(x)$ .

For the equality, assume that  $q \in \partial\varphi(x)$ , i.e. for all  $y$

$$f(Ax) + \langle q, y - x \rangle \leq f(Ay).$$

We aim to squeeze a separating hyperplane into this inequality.

We define

$$C_1 = \text{epi } f, \quad C_2 = \{(Ay, f(Ax) + \langle q, y - x \rangle) \in \mathbb{R}^d \times \mathbb{R} \mid y \in \mathbb{R}^d\}.$$

We note that  $\text{int}(C_1)$  is not empty and that  $\text{int}(C_1) \cap C_2 = \emptyset$ .

Hence, by Theorem 3.9 we can find some  $0 \neq (q^0, \alpha_0) \in \mathbb{R}^d \times \mathbb{R}$  such that

$$\begin{aligned}\langle q^0, \bar{y} \rangle + \alpha_0 \alpha &\leq \lambda \quad \forall \bar{y} \in \text{dom } f, \alpha \geq f(\bar{y}) \\ \langle q^0, Ay \rangle + \alpha_0 (f(Ax) + \langle q, y - x \rangle) &\geq \lambda, \quad \forall y \in \mathbb{R}^d.\end{aligned}$$

Again  $\alpha_0 > 0$  can not occur ( $\alpha \rightarrow \infty$  would lead to a contradiction) and  $\alpha_0 \neq 0$  follows from the continuity of  $f$  at  $A\bar{x}$ .

If  $\alpha_0 = 0$ , then  $\langle q^0, \bar{y} \rangle \leq \langle q^0, Ay \rangle$  for all  $\bar{y} \in \text{dom } f$  and  $y \in \mathbb{R}^d$ . Hence we can choose  $y = x \pm \Delta x$  and  $\bar{y} = A(x \pm \Delta x)$  and would get  $q^0 = 0$  as well, contradicting  $(q^0, \alpha_0) \neq 0$ .

If we set  $\bar{y} = Ax$ ,  $\alpha = f(\bar{y})$  and  $y = x$  we conclude  $\lambda = \langle q^0, Ax \rangle + \alpha_0 f(Ax)$ . By the second inequality above we get

$$\langle q^0, Ay - Ax \rangle + \alpha_0 \langle q, y - x \rangle \geq 0 \quad \forall y \in \mathbb{R}^d$$

and hence  $q = -\frac{1}{\alpha_0} A^T q^0$ . Setting  $p = -\frac{1}{\alpha_0} q^0$  we get from the first inequality above, by rearranging and dividing by  $\alpha_0 < 0$ , that

$$\langle p, z - Ax \rangle + f(Ax) \leq f(z), \quad \forall z \in \text{dom } f$$

and this means that  $p \in \partial f(Ax)$ . Hence,  $\partial\varphi(x) \subset A^T \partial f(Ax)$ .  $\square$

## 11 Applications of subgradients

*Example 11.1.* Let's consider counterexamples to show that the sum-rule rule and the chain rule for linear maps are indeed not always true:

1. Let  $f = i_{[0, \infty[}$  and

$$g(x) = \begin{cases} -\sqrt{-x}, & x \leq 0 \\ \infty, & x > 0. \end{cases}$$

The subdifferentials are

$$\partial f(x) = \begin{cases} \emptyset, & x < 0 \\ ]-\infty, 0], & x = 0 \\ \{0\}, & x > 0 \end{cases}, \quad \partial g(x) = \begin{cases} \{\frac{1}{2\sqrt{-x}}\}, & x < 0 \\ \emptyset, & x \geq 0. \end{cases}$$

Thus, the sum of the subdifferentials is always empty:  $\partial f(x) + \partial g(x) = \emptyset$  for all  $x$ .

The sum of  $f$  and  $g$ , however, is

$$(f + g)(x) = i_{\{0\}}(x)$$

and hence, the subdifferential of the sum is

$$\partial(f + g)(x) = \begin{cases} \emptyset, & x \neq 0, \\ \mathbb{R}, & x = 0, \end{cases}$$

which is much larger.

2. For the chain rule for linear maps there is a very simple counterexample: Consider

$$f(x) = \begin{cases} \infty, & \text{if } x < 0, \\ -\sqrt{x}, & \text{if } x \geq 0, \end{cases}$$

with  $\partial f(0) = \emptyset$  and  $A = 0$  (the  $1 \times 1$  zero-matrix). Then  $\varphi(x) = f(Ax) \equiv f(0) = 0$ , i.e.  $\partial \varphi(0) = \{0\}$ . Hence

$$\partial \varphi(0) = \{0\} \supsetneq \emptyset = A\partial f(A0).$$

△

Here is a positive result:

**Corollary 11.2.** Let  $f : \mathbb{R}^d \rightarrow \bar{\mathbb{R}}$  be proper, lsc, and convex and let  $C \subset \mathbb{R}^d$  be non-empty, closed, and convex such that  $\text{dom } f \cap \text{int } C \neq \emptyset$ . Then it holds

$$\hat{x} \in \underset{x \in C}{\operatorname{argmin}} f(x) \iff \begin{cases} 1. \hat{x} \in C \\ 2. 0 \in \partial f(\hat{x}) + N_C(\hat{x}) \end{cases}$$

The condition that  $\text{dom } f \cap \text{int } C \neq \emptyset$  ensures that the sub-gradient sum-rule is fulfilled for  $f + i_C$  and  $N_C(\hat{x}) = \partial i_C(\hat{x})$  proves the claim.

Here are two reformulations of the optimality conditions:  $\hat{x}$  solves  $\min_{x \in C} f(x)$  if and only if  $\hat{x} \in C$  and one of the following conditions holds

- A.  $\exists p \in \partial f(\hat{x}) \forall x \in C: \langle p, x - \hat{x} \rangle \geq 0$ ,
- B.  $\exists p \in \partial f(\hat{x}) \forall \gamma > 0: \hat{x} = P_C(\hat{x} - \gamma p)$ .

Condition A: It holds that  $0 \in \partial f(\hat{x}) + N_C(\hat{x})$  if and only if there is  $p \in \partial f(\hat{x})$  and  $-p \in N_C(\hat{x})$ . Writing out the latter condition gives the assertion.

Condition B: The Projection Theorem (Theorem 2.9) we can characterize  $\hat{x} = P_C(\hat{x} - \gamma p)$  by

$$\forall x \in C: \langle \hat{x} - \gamma p - \hat{x}, x - \hat{x} \rangle \leq 0$$

and this is equivalent to

$$\forall x \in C: \langle p, x - \hat{x} \rangle \geq 0.$$

The above results allow us to obtain very simple algorithms in several situations.

*Example 11.3* (Non-negative least squares). We consider a least squares problem  $\min \frac{1}{2} \|Ax - b\|_2^2$  and want to find only non-negative solutions, i.e. we add the constraint  $x \geq 0$ . We can do this by adding an indicator function of the non-negative orthant  $C = \mathbb{R}_{\geq 0}^d$ , i.e. we consider

$$\min_x \frac{1}{2} \|Ax - b\|_2^2 + i_{\mathbb{R}_{\geq 0}^d}(x).$$

We use the projection characterization (item B. above): Projecting onto the non-negative orthant is just clipping away the negative entries, i.e. we take the positive part of the vector:

$$P_{\mathbb{R}_{\geq 0}^d}(x) = \max(x, 0) =: x_+.$$

Moreover, the function  $f(x) = \frac{1}{2} \|Ax - b\|_2^2$  is convex and differentiable, hence  $\partial f(x) = \{A^T(Ax - b)\}$  and thus, solutions are characterized by

$$\hat{x} = (\hat{x} - \gamma A^T(A\hat{x} - b))_+$$

for any  $\gamma > 0$ . It turns out that for suitable  $\gamma$  (namely  $0 < \gamma < 2/\|A^T A\|$ ) the corresponding fixed point iteration

$$x^{k+1} = (x^k - \gamma A^T(Ax^k - b))_+$$

converges to a solution (we will prove this later). △

A little bit more general:

*Example 11.4* (Projected subgradient method). Let  $f : \mathbb{R}^d \rightarrow \bar{\mathbb{R}}$  be proper, convex and lsc and let  $C$  be non-empty, closed and convex and consider the general convexly constrained convex optimization problem  $\min_{x \in C} f(x)$

Condition B. now reads as: For some  $\hat{p} \in \partial f(\hat{x})$  it holds that

$$\hat{x} = P_C(\hat{x} - \gamma \hat{p})$$

which we can turn into a fixed-point iteration

$$\begin{aligned} \text{Choose } p^k &\in \partial f(x^k), \\ \text{Set } x^{k+1} &= P_C(x^k - \gamma p^k). \end{aligned}$$

Let us analyze this method a little bit: If  $x^*$  denotes any solution of the problem, then  $P_C(x^*) = x^*$  and since  $P_C$  is Lipschitz continuous with constant 1 (we will show this later) we get

$$\begin{aligned} \|x^{k+1} - x^*\|_2^2 &= \|P_C(x^k - \gamma p^k) - P_C(x^*)\|_2^2 \\ &\leq \|x^k - \gamma p^k - x^*\|_2^2 \\ &= \|x^k - x^*\|_2^2 - 2\gamma \langle p^k, x^k - x^* \rangle + \gamma^2 \|p^k\|_2^2 \\ &\leq \|x^k - x^*\|_2^2 - 2\gamma(f(x^k) - f(x^*)) + \gamma^2 \|p^k\|_2^2 \end{aligned}$$

where the last step uses the subgradient inequality  $f(x^*) \geq f(x^k) + \langle p^k, x^* - x^k \rangle$ . Since  $f(x^k) - f(x^*) \geq 0$  ( $x^*$  is a minimizer) we see that a step with a small enough stepsize  $\gamma$  should reduce the distance to any minimizer. Of course, we can also use a stepsize  $\gamma_k$  that changes with each iteration, leading to an estimate

$$\|x^{k+1} - x^*\|_2^2 \leq \|x^k - x^*\|_2^2 - 2\gamma_k(f(x^k) - f(x^*)) + \gamma_k^2 \|p^k\|_2^2$$

We can rearrange this to

$$\gamma_k(f(x^k) - f(x^*)) \leq \frac{1}{2}\gamma_k^2 \|p^k\|_2^2 + \frac{1}{2}\|x^k - x^*\|_2^2 - \frac{1}{2}\|x^{k+1} - x^*\|_2^2.$$

Now we sum up these inequalities for  $k = 0, \dots, N$  and get

$$\begin{aligned} \sum_{k=0}^N \gamma_k(f(x^k) - f(x^*)) &\leq \frac{1}{2} \sum_{k=0}^N \gamma_k^2 \|p^k\|_2^2 + \frac{1}{2}\|x^0 - x^*\|_2^2 - \frac{1}{2}\|x^{N+1} - x^*\|_2^2 \\ &\leq \frac{1}{2} \sum_{k=0}^N \gamma_k^2 \|p^k\|_2^2 + \|x^0 - x^*\|_2^2. \end{aligned}$$

To get a convergent method, we assume that the norms of the subgradients are bounded, i.e. for all  $k$  we have  $\|p^k\|_2 \leq L$  for some  $L > 0$ . Furthermore we denote  $f^* := f(x^*)$ ,  $f_{\text{best}}^N := \min_{k=0, \dots, N} f(x^k)$  and  $D^2 = \frac{1}{2}\|x^0 - x^*\|_2^2$ . Then we get

$$\left( \sum_{k=0}^N \gamma_k \right) (f_{\text{best}}^N - f^*) \leq \frac{L^2}{2} \sum_{k=0}^N \gamma_k^2 + D^2.$$

Finally, this leads to

$$(f_{\text{best}}^N - f^*) \leq \frac{D^2 + \frac{L^2}{2} \sum_{k=0}^N \gamma_k^2}{\sum_{k=0}^N \gamma_k}.$$

This shows: The best function value converges towards the minimal one, if we assume that the stepsizes  $\gamma_k$  fulfill

$$\frac{\sum_{k=0}^N \gamma_k^2}{\sum_{k=0}^N \gamma_k} \xrightarrow{N \rightarrow \infty} 0.$$

This can be accomplished, for example, if  $\sum_{k=0}^{\infty} \gamma_k^2$  converges, while  $\sum_{k=1}^{\infty} \gamma_k$  diverges which is the case, for example, for  $\gamma_k = 1/(k+1)$ . However, other stepsizes also make sense, e.g. one can consider  $\gamma_k = 1/\sqrt{k+1}$ . More precisely, we get the estimates (by comparison with the respective integrals)

We use that for a function  $f$  that decreases on an interval  $[K-1, N+1]$  it holds that  $\int_K^{N+1} f(x) dx \leq \sum_{k=K}^N f(k) \leq \int_{K-1}^N f(x) dx$ .

$$\begin{aligned} \log(N+2) &= \int_0^{N+1} \frac{1}{x+1} dx \leq \sum_{k=0}^N \frac{1}{k+1} \leq 1 + \int_0^N \frac{1}{x+1} dx = 1 + \log(N+1) \\ 2\sqrt{N+2} - 2 &= \int_0^{N+1} \frac{1}{\sqrt{x+1}} dx \leq \sum_{k=0}^N \frac{1}{\sqrt{k+1}} \leq 1 + \int_0^N \frac{1}{\sqrt{x+1}} dx = 2\sqrt{N+1} - 1 \\ &\quad \sum_{k=0}^N \frac{1}{(k+1)^2} \leq 1 + \int_0^N \frac{1}{(x+1)^2} dx = 2 - \frac{1}{N+1}. \end{aligned}$$

Hence, we get for the stepsize  $\gamma_k = 1/(k+1)$  that

$$(f_{\text{best}}^N - f^*) \leq \frac{D^2 + \frac{L^2}{2} (2 - \frac{1}{N+1})}{\log(N+2)}.$$

and for the stepsize  $\gamma_k = 1/\sqrt{k+1}$  that

$$(f_{\text{best}}^N - f^*) \leq \frac{D^2 + \frac{L^2}{2} (1 + \log(N+1))}{2\sqrt{N+2} - 2}.$$

If one wants to achieve the best result with a fixed number  $N$  of iterations, one can proceed differently: Here one would like to minimize the ratio

$$R(\gamma) = \frac{D^2 + \frac{L^2}{2} \sum_{k=0}^N \gamma_k^2}{\sum_{k=0}^N \gamma_k}$$

over all variables  $\gamma_k$ . We can take the gradient with respect to the vector  $\gamma = (\gamma_1, \gamma_2, \dots, \gamma_N)$  and get

$$\nabla R(\gamma) = \frac{L^2 \gamma \sum_{k=0}^N \gamma_k - (D^2 + \frac{L^2}{2} \sum_{k=0}^N \gamma_k^2) \mathbf{1}}{\left( \sum_{k=0}^N \gamma_k \right)^2}$$

where  $\mathbb{1}$  denotes the vector of all ones. Setting this to zero we observe that  $\gamma$  should be a constant vector, i.e.  $\gamma = h\mathbb{1}$  for some  $h$ . Plugging this in, we need to solve

$$L^2(N+1)h^2 = \frac{L^2}{2}(N+1)h^2 + D^2$$

leading to the constant stepsize

$$\gamma_k \equiv \frac{\sqrt{2}D}{L\sqrt{N+1}}.$$

This gives us the guarantee that

$$f_{\text{best}}^N - f^* \leq 2D^2 \frac{1}{\sqrt{N+1}}.$$

In conclusion we get the following:

stepsize $\gamma_k$	$\frac{1}{k+1}$	$\frac{1}{\sqrt{k+1}}$	$\frac{\sqrt{2}D}{L\sqrt{N+1}}$
estimate $f_{\text{best}}^N - f^*$	$\frac{L^2}{2} \frac{D^2+2-\frac{1}{N+1}}{\log(N+2)}$	$\frac{L^2}{2} \frac{D^2+1+\log(N+1)}{2\sqrt{N+2}-2}$	$\frac{2D^2}{\sqrt{N+1}}$

So, theoretically, the option with fixed step-size has the best worst-case guarantee, but note that further iterations will not improve this any more, and moreover, an estimate on  $\|x^0 - x^*\|_2$  is needed. In practice, the stepsize  $1/(k+1)$  often leads to best results, but the choice of step-sizes for subgradient methods is a delicate issue.

Here is an example for the problem of *least absolute deviations* (LAD) which is

$$\min_{x \in \mathbb{R}^n} \|Ax - b\|_1,$$

i.e. one minimizes the sum of the absolute deviations and not the sum of the squares. This approach is useful if  $b$  contains additive noise which follows a Laplace distribution. Since the 1-norm is convex, lsc and everywhere continuous, we can apply the chain rule from Theorem 10.4. Since the subgradient of the 1-norm is obtained from the known subgradient of the absolute value by applying this component-wise, we obtain

$$\partial\|x\|_1 = \text{Sign}(x)$$

with the so-called *multivalued sign* function which acts component-wise as

$$\text{Sign}(x_i) = \begin{cases} \{-1\}, & x_i < 0 \\ [-1, 1], & x_i = 0 \\ \{1\}, & x_i > 0. \end{cases}$$

By the chain rule and Lemma 9.6 we get as subgradient of  $f(x) = \|Ax - b\|_1$

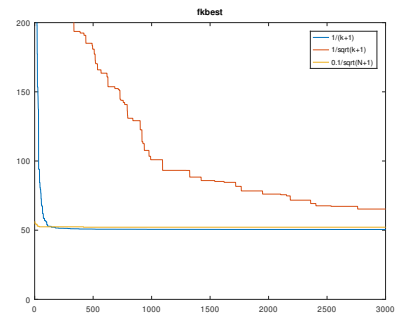
$$\partial f(x) = A^T \text{Sign}(Ax - b).$$



Hence, the subgradient method for the LAD problem can be implemented by choosing the ordinary sign-function

$$x^{k+1} = x^k - \gamma_k A^T \text{sign}(Ax^k - b).$$

$\Delta$



## 12 Proximal algorithms

**Lemma 12.1.** *The proximal mapping  $\text{prox}_{\lambda f}$  for a convex and lsc function and any  $\lambda > 0$  is Lipschitz continuous with constant 1, i.e. it holds*

$$\|\text{prox}_{\lambda f}(x) - \text{prox}_{\lambda f}(y)\|_2 \leq \|x - y\|_2.$$

*Proof.* We use the variational inequality from Theorem 8.2 for  $\hat{x} = \text{prox}_{\lambda f}(x)$  and  $\hat{y} = \text{prox}_{\lambda f}(y)$  and plug in  $\hat{y}$  and  $\hat{x}$  is the other inequality to get

$$\begin{aligned}\langle x - \hat{x}, \hat{y} - \hat{x} \rangle + \lambda(f(\hat{x}) - f(\hat{y})) &\leq 0 \\ \langle y - \hat{y}, \hat{x} - \hat{y} \rangle + \lambda(f(\hat{y}) - f(\hat{x})) &\leq 0.\end{aligned}$$

Adding these inequalities gives

$$\begin{aligned}0 &\geq \langle x - \hat{x}, \hat{y} - \hat{x} \rangle + \langle y - \hat{y}, \hat{x} - \hat{y} \rangle \\ &= \langle x - y - (\hat{x} - \hat{y}), \hat{y} - \hat{x} \rangle \\ &= \langle x - y, \hat{y} - \hat{x} \rangle + \|\hat{y} - \hat{x}\|_2^2.\end{aligned}$$

By Cauchy-Schwarz, we get

$$\|\hat{y} - \hat{x}\|_2^2 \leq \|x - y\|_2 \|\hat{y} - \hat{x}\|_2$$

which shows the claim.  $\square$

We start with a simple algorithm to solve an unconstrained minimization problem  $\min_x f(x)$  (the algorithm is, in this very simple form, not practically useful, but will be good to know, since it can be used as a building block for further methods). The method is called the *proximal point method* and simply iterates

$$x^{k+1} = \text{prox}_{t_k f}(x^k)$$

for some sequence  $t_k > 0$  of stepsizes.

**Lemma 12.2.** *The sequence  $(x^k)$  from the proximal point method fulfills*

$$f(x^{k+1}) \leq f(x^k) - \frac{1}{t_k} \|x^{k+1} - x^k\|_2^2.$$

*Proof.* From the variational inequality (Theorem 8.2) we get with  $x = y = x^k$  and  $\hat{x} = x^{k+1}$  that

$$t_{k+1}(f(x^{k+1}) - f(x^k)) + \langle x^k - x^{k+1}, x^k - x^{k+1} \rangle \leq 0$$

from which the claim follows.  $\square$

This can be used to show convergence of the method:

**Theorem 12.3.** *Let  $f$  be proper, convex and lsc and let  $f^* = f(x^*) = \min_x f(x)$ . Then it holds that the sequence  $(x^k)$  generated by the proximal point method fulfills*

$$\begin{aligned}\|x^{k+1} - x^*\|_2 &\leq \|x^k - x^*\|_2 \text{ and} \\ f(x^{N+1}) - f^* &\leq \frac{\|x^0 - x^*\|_2^2}{2 \sum_{k=0}^N t_k}.\end{aligned}$$

*Proof.* Again by the variational inequality from Theorem 8.2 (now with  $x = x^k$ ,  $\hat{x} = x^{k+1}$  and  $y = x^*$ )

$$\begin{aligned} f(x^{k+1}) &\leq f^* - \frac{1}{t_k} \langle x^k - x^{k+1}, x^* - x^{k+1} \rangle \\ &= f^* + \frac{1}{t_k} \left( \langle x^k - x^{k+1}, x^k - x^* \rangle - \|x^k - x^{k+1}\|_2^2 \right) \end{aligned}$$

where we inserted  $-x^k + x^k$  to get to the second equality. Now we use  $\langle a, b \rangle = \frac{1}{2}(\|a\|_2^2 + \|b\|_2^2 - \|a - b\|_2^2)$  with  $a = x^k - x^{k+1}$  and  $b = x^k - x^*$  to get

$$\begin{aligned} f(x^{k+1}) &\leq f^* + \frac{1}{2t_k} \left( \|x^k - x^{k+1}\|_2^2 + \|x^k - x^*\|_2^2 - \|x^{k+1} - x^*\|_2^2 - 2\|x^k - x^{k+1}\|_2^2 \right) \\ &\leq f^* + \frac{1}{2t_k} \left( \|x^k - x^*\|_2^2 - \|x^{k+1} - x^*\|_2^2 \right). \end{aligned}$$

Since  $f^* \leq f(x^{k+1})$  we get, on the one hand,  $\|x^{k+1} - x^*\|_2 \leq \|x^k - x^*\|_2$  and, on the other hand, summing up the estimates from  $k = 0, \dots, N$  we get

$$2 \sum_{k=0}^N t_k (f(x_{k+1}) - f^*) \leq \|x^0 - x^*\|_2^2.$$

Since we already know that  $f(x^{k+1}) - f^* \geq f(x^{N+1}) - f^*$  we get

$$2 \left( \sum_{k=0}^N t_k \right) (f(x^{N+1}) - f^*) \leq \|x^0 - x^*\|_2^2$$

as claimed.  $\square$

Hence, we see that larger stepsizes are better, but one should emphasize that the proximal point method is merely a theoretical algorithm, as each step needs the evaluation of the proximal map for the objective and this may be no simpler than the original problem.

Now we come to a more practical algorithm and this relies on a very fruitful idea: If the objective function in our optimization problem is the sum of two convex function, i.e.

In fact it can be simpler, since the objective in the minimization problem for the prox is strongly convex even if  $f$  is just convex.

$$\min_{x \in \mathbb{R}^d} f(x) + g(x)$$

we may try to treat both terms differently, depending on their properties. Methods that are derived from splitting the objective additively into different parts go under the name *splitting methods*. Two different properties that will be useful are the following:

- *L-smoothness*:  $g : \mathbb{R}^d \rightarrow \mathbb{R}$  is convex and differentiable and, moreover, that the gradient  $\nabla g$  is Lipschitz-continuous with constant  $L$ , i.e.  $\|\nabla g(y) - \nabla g(x)\|_2 \leq L\|y - x\|_2$ .

- *proximability*:  $f$  is of the form that  $\text{prox}_{\lambda f}$  is simple to evaluate. This is the case, for example for the 1-norm  $\|x\|_1$  or indicator functions  $i_C$  of convex sets if the projection onto these sets is simple (e.g. positivity constraints, hyperplanes, balls).

We have just seen in Lemma 12.2 that the proximal map reduces the objective (for this part) and the next lemma shows that one can also get a guaranteed descent of the objective by doing gradient steps for  $L$ -smooth functions:

**Lemma 12.4.** *If  $g : \mathbb{R}^d \rightarrow \mathbb{R}$  is  $L$ -smooth, then it holds that*

$$g(y) - g(x) - \langle \nabla g(x), y - x \rangle \leq \frac{L}{2} \|x - y\|_2^2.$$

*Proof.* By the fundamental theorem of calculus we get

$$\begin{aligned} g(y) - g(x) - \langle \nabla g(x), y - x \rangle &= \int_0^1 \langle \nabla g(x + \tau(y - x)) - \nabla g(x), y - x \rangle d\tau \\ &\leq \int_0^1 \|\nabla g(x + \tau(y - x)) - \nabla g(x)\|_2 \|y - x\|_2 d\tau \\ &\leq \int_0^1 L \tau \|y - x\|_2 d\tau \\ &\leq \frac{L}{2} \|x - y\|_2^2. \end{aligned}$$

□

This allows to guarantee a reduction the value of  $g(x)$  by making a gradient step  $x^+ = x - \lambda \nabla g(x)$ : We use the above lemma with  $y = x^+$  and the fact that  $x^+ - x = -\lambda \nabla g(x)$  to get

$$\begin{aligned} g(x^+) &\leq g(x) + \langle \nabla g(x), x^+ - x \rangle + \frac{L}{2} \|x^+ - x\|_2^2 \\ &= g(x) - \lambda \|\nabla g(x)\|_2^2 + \frac{L\lambda^2}{2} \|\nabla g(x)\|_2^2 \\ &= g(x) - \lambda(1 - \frac{L}{2}\lambda) \|\nabla g(x)\|_2^2. \end{aligned}$$

Hence, we get a guaranteed descent if both  $\lambda > 0$  and  $1 - \frac{L}{2}\lambda > 0$ , i.e. if  $0 < \lambda < 2/L$ .

How can we use these two ingredients to come up with a method that can minimize the objective  $f + g$ ? One idea is, to replace the differentiable part of the objective by a simpler upper bound at some current iterate  $x^k$ : Inspired by Lemma 12.4 we define  $x^{k+1}$  by

$$x^{k+1} = \underset{x}{\operatorname{argmin}} \left[ f(x) + g(x^k) + \langle \nabla g(x^k), x - x^k \rangle + \frac{1}{2\lambda} \|x - x^k\|_2^2 \right].$$

We can drop the terms in the objective which do not depend on  $x$  and multiply by  $\lambda$  to get

$$x^{k+1} = \underset{x}{\operatorname{argmin}} \left[ \lambda f(x) + \langle \lambda \nabla g(x^k), x - x^k \rangle + \frac{1}{2\lambda} \|x - x^k\|_2^2 \right].$$

We complete the square to see that

$$\begin{aligned} & \langle \lambda \nabla g(x^k), x - x^k \rangle + \frac{1}{2\lambda} \|x - x^k\|_2^2 \\ &= \frac{1}{2} \|\lambda \nabla g(x^k)\|_2^2 + \langle \lambda \nabla g(x^k), x - x^k \rangle + \frac{1}{2\lambda} \|x - x^k\|_2^2 - \frac{1}{2} \|\lambda \nabla g(x^k)\|_2^2 \\ &= \frac{1}{2} \|x - (x^k - \lambda \nabla g(x^k))\|_2^2 - \frac{1}{2} \|\lambda \nabla g(x^k)\|_2^2. \end{aligned}$$

Again dropping terms that do not affect the minimizer, we finally get that

$$\begin{aligned} x^{k+1} &= \operatorname{argmin}_x \lambda f(x) + \frac{1}{2} \|x - (x^k - \lambda \nabla g(x^k))\|_2^2 \\ &= \operatorname{prox}_{\lambda f}(x^k - \lambda \nabla g(x^k)). \end{aligned}$$

This method does a gradient step for the smooth part and a proximal step for the proximable part.

*Example 12.5* (Regularized least squares for inverse problems). One example for which the proximal gradient method gained popularity is the case of regularized least squares problems: For some  $A \in \mathbb{R}^{m \times n}$  and  $b \in \mathbb{R}^m$  one can consider the least squares problem  $\min_x \frac{1}{2} \|Ax - b\|_2^2$ . This is used in many contexts, e.g. in statistics for regression but also in the context of signal processing or inverse problems where some quantity of interest  $x^\dagger \in \mathbb{R}^n$  can only be measured indirectly, namely one can only observe  $b^\delta = Ax^\dagger + \eta$  where  $A$  is a (known) linear map which models the measurement process, and  $\eta$  is an unknown error (e.g. due to measurement noise or modelling errors). Since  $b^\delta$  is also affected by noise, it is pointless to solve  $Ax = b^\delta$  exactly and a least squares approach seems more reasonable. In addition it may also happen that  $n > m$ , i.e. we do not have enough measurements to reconstruct  $x^\dagger$  even from a noise-free  $b^\delta$ . More precisely, the minimizers of the least squares problem are characterized by the equation  $A^T Ax = A^T b^\delta$ , but since  $A^T A$  does not have full rank, there are still multiple solutions, but due to noise, none of these seems to be reasonable. In this case one uses prior knowledge on the unknown solution, and this is done by specifying a *regularization functional*  $R : \mathbb{R}^n \rightarrow \bar{\mathbb{R}}$  which gives small values  $R(x)$  to “reasonable”  $x$  and large values  $R(x)$  for “undesired”  $x$ . The regularization functional is also called *penalty function* since it penalizes undesired vectors. Together, we end with the *regularized least squares problem*

$$\min_x \frac{1}{2} \|Ax - b^\delta\|_2^2 + \alpha R(x)$$

where  $\alpha$  is a positive *regularization parameter* that can emphasize the regularization (large  $\alpha$ ) or tone it down (small  $\alpha$ ).

If  $R$  is convex, lsc and proximable, one can use the proximal gradient method to solve this minimization problem: We take  $g(x) = \frac{1}{2} \|Ax - b^\delta\|_2^2$  and  $f(x) = \alpha R(x)$  and since  $\nabla g(x) = A^T(Ax - b^\delta)$  is Lipschitz continuous with constant  $L = \|A^T A\|$  one iterates

$$x^{k+1} = \operatorname{prox}_{\lambda \alpha R}(x^k - \lambda A^T(Ax^k - b^\delta))$$

with some  $\lambda \in ]0, 2/\|A^T A\|]$ .

$\triangle$

### 13 Convex conjugation

There is an important notion of duality for convex functions, namely, the one of convex conjugation (and this duality is related to the characterization of closed, convex set as intersection of half-spaces). In Corollary 5.5 we have already seen that a closed convex function is equal to the supremum of all affine functions that are below said function. This leads to the following definition:

**Definition 13.1.** For a function  $f : \mathbb{R}^d \rightarrow \bar{\mathbb{R}}$  we define the (Fenchel) conjugate  $f^* : \mathbb{R}^d \rightarrow \bar{\mathbb{R}}$  as

$$f^*(p) = \sup_x [\langle p, x \rangle - f(x)]$$

Moreover, the biconjugate is

$$f^{**}(x) = \sup_p [\langle p, x \rangle - f^*(p)].$$

If  $f$  is proper, then  $f^*$  is the pointwise supremum of affine functions and hence, it is always convex and lower semicontinuous (even when  $f$  has none of these properties).

We consider a few one-dimensional examples:

**Example 13.2.** 1. Let  $f(x) = \alpha x^2$  for  $\alpha > 0$ . Then the conjugate is

$$f^*(p) = \sup_x [px - \alpha x^2].$$

To calculate the supremum we take the derivative with respect to  $x$ , set it to zero and plug it in:

$$\begin{aligned} p - 2\alpha x &= 0 \implies x = \frac{p}{2\alpha} \\ \implies \sup_x [px - \alpha x^2] &= p \frac{p}{2\alpha} - \alpha \frac{p^2}{4\alpha^2} = \frac{1}{4\alpha} p^2, \end{aligned}$$

hence,

$$f^*(p) = \frac{1}{4\alpha} p^2.$$

Similarly, one observes that  $f^{**}(x) = f(x)$ . (Note that  $f^* = f$  for  $\alpha = \frac{1}{2}$ , i.e. for  $f(x) = x^2/2$ .)

2. Let  $f(x) = \exp(x)$ . We consider different cases:

- If  $p < 0$ , then  $px - \exp(x)$ , is unbounded from above and we get  $f^*(p) = \infty$ .
- If  $p = 0$ , then  $-\exp$  has the supremum 0 (which is not attained) and we have  $f^*(0) = 0$ .

- If  $p > 0$ , then  $px - \exp(x)$  is bounded from above and we calculate as in the previous example

$$\begin{aligned} p - \exp(x) = 0 &\implies x = \log(p) \\ &\implies \sup_x [px - \exp(x)] = p \log(p) - p. \end{aligned}$$

Hence, the conjugate is

$$f^*(p) = \begin{cases} \infty, & \text{if } p < 0 \\ 0, & \text{if } p = 0 \\ p \log(p) - p, & \text{if } p > 0. \end{cases}$$

One may verify in a similar way that  $f^{**}(x) = \exp(x) = f(x)$ .

3. For  $f(x) = |x|$  we have  $f^*(x) = \sup_x px - |x|$  and we see by direct inspection that

$$f^*(x) = \begin{cases} \infty, & \text{if } p > 1 \\ 0, & \text{if } -1 \leq p \leq 1 \\ \infty, & \text{if } p < -1 \end{cases}$$

i.e.  $f^*(x) = i_{[-1,1]}(x)$ . Again, we get  $f^{**}(x) = |x| = f(x)$ .

4. For a non-convex example, consider

$$f(x) = \begin{cases} \infty, & \text{if } |x| > 1 \\ 1 - x^2, & \text{if } |x| \leq 1. \end{cases}$$

In this case one has  $f^*(x) = |x|$  and by the previous example,  $f^{**}(x) = i_{[-1,1]}(x) \neq f(x)$ . However,  $f^{**} = \text{cl } f$ .

△

We will see later that the observed behavior  $f^{**} = f$  for convex and lsc functions and  $f^{**} = \text{cl } f$  are true in general.

**Lemma 13.3.** Let  $f, g : \mathbb{R}^d \rightarrow \bar{\mathbb{R}}$ . Then:

- i) If  $f \geq g$ , then  $f^* \leq g^*$ .
- ii) For all  $p, x$  where  $f(x)$  or  $f^*(p)$  are finite we have Fenchel's inequality

$$f(x) + f^*(p) \geq \langle x, p \rangle.$$

- iii) It holds Fenchel's equality

$$p \in \partial f(x) \iff f(x) + f^*(p) = \langle p, x \rangle \iff x \in \partial f^*(p).$$

The equivalence  $p \in \partial f(x) \iff x \in \partial f^*(p)$  is called *subgradient inversion theorem*.

- iv) It holds  $f^{**} \leq f$ .



*Proof.* i) If  $f \geq g$ , then

$$\begin{aligned} g^*(p) &= \sup_x [\langle p, x \rangle - g(x)] \\ &\geq \sup_x [\langle p, x \rangle - f(x)] = f^*(p). \end{aligned}$$

ii) Follows directly from the definition of the conjugate by replacing the supremum by any value of the argument.

iii) By rearranging the subgradient inequality, we have  $p \in \partial f(x)$  if and only if for all  $y$  it holds  $\langle x, p \rangle - f(x) \geq \langle p, y \rangle - f(y)$ . Taking the supremum over all  $y$  shows  $\langle x, p \rangle - f(x) \geq f^*(p)$  and by Fenchel's inequality, we get equality. Since this argument works both ways, the claim is proven.

iv) We consider

$$f^{**}(x) = \sup_p [\langle p, x \rangle - f^*(p)]$$

and use ii) to estimate the term in the supremum by  $f(x)$  (which is then independent from  $p$ ).

□

**Lemma 13.4.** Let  $f : \mathbb{R}^d \rightarrow \bar{\mathbb{R}}$ . Then it holds:

- i) If  $f$  is not proper, then  $f^* \equiv \infty$  or  $\equiv -\infty$ .
- ii) If  $f$  is proper, then  $f^* > -\infty$ .
- iii) If  $f$  is proper and convex, then  $f^*$  is proper and moreover we have  $(\text{cl } f)^* = f^*$  and  $f^{**} = \text{cl } f$ .

*Proof.* i) For non-proper  $f$  we have two cases. In the first case there is  $x_0$  such that  $f(x_0) = -\infty$ . But then there is no affine function below  $f$  and hence  $f^* \equiv \infty$ . In the second case  $f \equiv \infty$ , end then  $f^* \equiv -\infty$ .

ii) Now let  $f$  be proper. If we had  $f^*(p_0) = -\infty$  for some  $p_0$ , then we would have for every  $x$  that  $-\infty \geq \langle p_0, x \rangle - f(x)$ , but this implies  $f(x) \geq \infty$  for all  $x$ .

- iii) • We show that  $f^*$  is proper: By Proposition 5.4, we know that there exist  $p_0 \in \mathbb{R}^d$  and  $\alpha \in \mathbb{R}$  such that  $f(x) \geq \langle p_0, x \rangle + \alpha$  for all  $x$ , and hence  $f^*(p_0) \leq -\alpha$ .
- Since  $f$  is proper, we have to show  $(\bar{f})^* = f^*$  and by Lemma 13.3 i) we deduce from  $\bar{f} \leq f$  the inequality  $(\bar{f})^* \geq f^*$ . For the lsc envelope we have by the Fenchel inequality that

$$\bar{f}(x_0) = \liminf_{x \rightarrow x_0} f(x) \geq \liminf_{x \rightarrow x_0} [\langle p, x \rangle - f^*(p)] = \langle p, x_0 \rangle - f^*(p).$$

This leads to  $f^*(p) \geq \langle p, x_0 \rangle - \bar{f}(x_0)$  and taking the supremum over all  $x_0$  shows  $f^* \geq (\bar{f})^*$ .

The statement  $f^{**} = \text{cl } f$  follows from the next theorem.

The theorem says that  $f = f^{**}$  if and only if  $f$  is convex and lsc, and hence  $f^{**}(f^*)^* = ((\text{cl } f)^*)^* = \text{cl } f$  since  $\text{cl } f$  is convex and lsc.

□

**Theorem 13.5** (Fenchel-Moreau). A proper function  $f : \mathbb{R}^d \rightarrow \mathbb{R}^d$  is convex and lsc if and only if  $f = f^{**}$

*Proof.* Since any conjugate is always convex and lsc, the reverse implication is clear.

Now let  $f$  be proper, convex and lsc. We already know that  $f^*$  is proper, convex and lsc as well and that  $f^{**} \leq f$ . Now we show the following claim:

If  $f(x) > \alpha$  for some  $x$  and  $\alpha$ , then  $f^{**}(x) \geq \alpha$ .

Once, the claim is proven, the theorem follows: If  $x \notin \text{dom}(f)$ , then we can choose  $\alpha$  arbitrarily large and see that  $f^{**}(x) = \infty$  as well. If  $x \in \text{dom}(f)$ , and  $f^{**}(x) = f(x) - \epsilon$  for some  $\epsilon > 0$ , then we could choose  $\alpha = f(x) - \epsilon/2$  and would get  $f^{**}(x) < \alpha$  which would contradict  $f^{**}(x) \geq \alpha$ .

Now we prove the claim: If  $(x, \alpha) \notin \text{epi}(f)$ , we can, since  $\text{epi}(f)$  is closed and convex, strictly separate  $(x, \alpha)$  from  $\text{epi}(f)$ , i.e. there is  $(p, a) \in \mathbb{R}^{d+1}$  and  $\epsilon > 0$  such that

$$\sup_{(y, \beta) \in \text{epi}(f)} \left\langle \begin{bmatrix} p \\ a \end{bmatrix}, \begin{bmatrix} y \\ \beta \end{bmatrix} \right\rangle \leq \left\langle \begin{bmatrix} p \\ a \end{bmatrix}, \begin{bmatrix} x \\ \alpha \end{bmatrix} \right\rangle - \epsilon.$$

which says that for all  $(y, \beta) \in \text{epi}(f)$  we have

$$\langle p, y \rangle + a\beta \leq \langle p, x \rangle + a\alpha - \epsilon \quad (*)$$

If we had  $a > 0$  we would get a contradiction with  $\beta \rightarrow \infty$ .

Hence we have  $a \leq 0$ . If  $a < 0$ , we divide  $(*)$  by  $-a > 0$  and set  $\bar{p} = -p/a$  to get

$$\langle \bar{p}, y \rangle - \beta \leq \langle \bar{p}, x \rangle - \alpha + \frac{\epsilon}{a}$$

We set  $\beta = f(y)$  and take the supremum over all  $y$  on the left-hand side to get

$$f^*(\bar{p}) \leq \langle \bar{p}, x \rangle - \alpha + \frac{\epsilon}{a} < \langle \bar{p}, x \rangle - \alpha.$$

We rearrange and use the Fenchel inequality (Lemma 13.3) to get

$$\alpha < \langle \bar{p}, x \rangle - f^*(\bar{p}) \leq f^{**}(x).$$

In the remaining case  $a = 0$  the inequality  $(*)$  turns into

$$\langle p, y \rangle \leq \langle p, x \rangle - \epsilon$$

and still holds for all  $y \in \text{dom}(f)$ . If we had  $x \in \text{dom}(f)$  we would get a contradiction by choosing  $y = x$ , so we have  $x \notin \text{dom}(f)$  if

$a = 0$ . So we have  $f(x) = \infty$  in this case and for all  $q$  we have by Fenchel's inequality again

$$\langle q, y \rangle - f(y) \leq f^*(q).$$

Multiplying the second to last inequality by  $\mu > 0$  and adding the last one gives

$$\langle q + \mu p, y \rangle - f(y) \leq f^*(q) + \mu \langle p, x \rangle - \mu \epsilon.$$

Taking the supremum over  $y$  gives

$$f^*(q + \mu p) \leq f^*(q) + \mu \langle p, x \rangle - \mu \epsilon,$$

which leads to

$$\langle q, x \rangle - f^*(q) + \mu \epsilon \leq \langle q + \mu p, x \rangle - f^*(q + \mu p) \leq f^{**}(x).$$

Since the left hand side goes to  $\infty$  for  $\mu \rightarrow \infty$ , this shows  $f^{**}(x) = \infty$ , as desired.  $\square$

## 14 Conjugation calculus

**Proposition 14.1.** If  $f : \mathbb{R}^d \rightarrow \bar{\mathbb{R}}$  is proper, then  $f^* = f$  if and only if  $f(x) = \frac{1}{2}\|x\|_2^2$ .

*Proof.* We have seen in the previous lecture that  $ax^2$  has the conjugate  $\frac{1}{4a}p^2$  so with  $a = \frac{1}{2}$  they are equal. Applying this component-wise we see that the conjugate of  $\frac{1}{2}\|x\|_2^2$  is the function itself.

For the converse, assume  $f = f^*$ . By Fenchel's inequality we have  $f(x) + f^*(p) \geq \langle x, p \rangle$  and with  $p = x$  we get  $f(x) \geq \frac{1}{2}\|x\|_2^2 =: h(x)$ . By Lemma 13.3 i) we know that  $f(x) = f^*(x) \leq h^*(x) = \frac{1}{2}\|x\|_2^2$  which proves the claim.  $\square$

**Lemma 14.2.** Let  $f : \mathbb{R}^d \rightarrow \bar{\mathbb{R}}$ ,  $A \in \mathbb{R}^{d \times d}$  be invertible,  $a \in \mathbb{R}$ ,  $b \in \mathbb{R}^d$  and  $\lambda \in \mathbb{R} \setminus \{0\}$ . Then it holds that

- i) If  $\varphi(x) = f(x) + a$ , then  $\varphi^*(p) = f^*(p) - a$ .
- ii) If  $\varphi(x) = f(\lambda x)$ , then  $\varphi^*(p) = f^*(\lambda^{-1}p)$ .
- iii) If  $\varphi(x) = \lambda f(x)$ ,  $\lambda > 0$ , then  $\varphi^*(p) = \lambda f^*(\lambda^{-1}p)$ .
- iv) If  $\varphi(x) = f(x) - \langle b, x \rangle$ , then  $\varphi^*(p) = f^*(p + b)$ .
- v) If  $\varphi(x) = f(Ax + b)$ , then  $\varphi^*(p) = f^*(A^{-T}p) - \langle A^{-T}p, b \rangle$ .

*Proof.* These are straightforward calculations:

- i)  $\varphi^*(p) = \sup_x [\langle p, x \rangle - f(x) - a] = f^*(p) - a$ .
- ii) This is a special case of v).
- iii)  $\varphi^*(p) = \sup_x [\langle p, x \rangle - \lambda f(x)] = \lambda \sup_x [\langle \lambda^{-1}p, x \rangle - f(x)] = \lambda f^*(\lambda^{-1}p)$ .
- iv)  $\varphi^*(p) = \sup_x [\langle p, x \rangle - (f(x) - \langle b, x \rangle)] = \sup_x [\langle p + b, x \rangle - f(x)] = f^*(p + b)$ .
- v)  $\varphi^*(p) = \sup_x [\langle p, x \rangle - f(Ax + b)] = \sup_y [\langle p, A^{-1}(y - b) \rangle - f(y)] = \sup_y [\langle A^{-T}p, y - b \rangle - f(y)] = f^*(A^{-T}p) - \langle A^{-T}p, b \rangle$ .

$\square$

By these rules one immediately sees that for  $f(x) = \frac{1}{2}\|x - b\|_2^2$  one has  $f^*(p) = \frac{1}{2}\|p\|_2^2 + \langle p, b \rangle$ , for example.

The conjugation of the infimal convolution is readily calculated:

**Lemma 14.3.** If  $f_1, f_2 : \mathbb{R}^d \rightarrow \bar{\mathbb{R}}$  are proper, then it holds that

$$(f_1 \square f_2)^* = f_1^* + f_2^*.$$

*Proof.* This can be seen as follows:

$$\begin{aligned}
 (f_1 \square f_2)^*(p) &= \sup_x [\langle p, x \rangle - \inf_{x_1+x_2=x} (f_1(x_1) + f_2(x_2))] \\
 &= \sup_x \sup_{x_1+x_2=x} [\langle p, x \rangle - f_1(x_1) - f_2(x_2)] \\
 &= \sup_{x_1, x_2} [\langle p, x_1 \rangle + \langle p, x_2 \rangle - f_1(x_1) - f_2(x_2)] \\
 &= f_1^*(p) + f_2^*(p).
 \end{aligned}$$

□

One would be tempted to assume that one also has that  $(f_1 + f_2)^* = f_1^* \square f_2^*$  but this does not hold without further assumptions.

Here is a counterexample:

*Example 14.4.* Let  $f, g : \mathbb{R}^2 \rightarrow \bar{\mathbb{R}}$  defined by

$$f(x) = \begin{cases} -\sqrt{x_1 x_2}, & x_1, x_2 \geq 0 \\ \infty, & \text{else} \end{cases}, \quad g(x) = I_{\{x_1=0\}}(x).$$

The sum is  $(f + g)(x) = I_{\{x_1=0, x_2 \geq 0\}}(x)$  and has the conjugate

$$(f + g)^*(p) = \sup_{x_1=0, x_2 \geq 0} [p_1 x_1 + p_2 x_2] = I_{\{p_2 \leq 0\}}(p).$$

The individual conjugates are

$$\begin{aligned}
 g^*(p) &= \sup_{x_1=0} [p_1 x_1 + p_2 x_2] = I_{\{p_2=0\}}(p) \\
 f^*(p) &= \sup_{x_1, x_2 \geq 0} [p_1 x_1 + p_2 x_2 + \sqrt{x_1 x_2}]
 \end{aligned}$$

We see that

$$f^*(p) = I_{\{p_1 p_2 \leq -1\}}(p).$$

Hence, the infimal convolution is

$$(f^* \square g^*)(p) = \inf_{q+r=p} I_{\{p_2=0\}}(q) + I_{\{p_1 p_2 \leq -1\}}(r) = I_{\{p_2=0\} + \{p_1 p_2 \leq -1\}}(p) = I_{\{p_2 < 0\}}(p).$$

Note that  $(f + g)^* \neq (f^* \square g^*)(p)$  and that the latter is not even lsc. △

We show a very general result which is due to Attouch and Brezis from 1986: Let  $f, g : \mathbb{R}^d \rightarrow \bar{\mathbb{R}}$  be two proper and convex functions and assume that

$$\bigcup_{\lambda \geq 0} \lambda(\text{dom}(f) - \text{dom}(g)) \text{ is a subspace of } \mathbb{R}^d. \quad (\text{Q})$$

Note that in Example 14.4 we have  $\text{dom}(f) = \{x_1, x_2 \geq 0\}$  and  $\text{dom}(g) = \{x_2 = 0\}$  and hence  $\text{dom}(f) - \text{dom}(g) = \{x_2 \geq 0\}$ . Thus, (Q) is not fulfilled

**Theorem 14.5.** If  $f, g : \mathbb{R}^d \rightarrow \bar{\mathbb{R}}$  are proper, convex, and lsc and satisfy condition (Q), then

$$(f + g)^* = f^* \square g^*$$

and the inf-convolution on the right is exact.

*Proof.* Step 0: First we note that for any decomposition  $p = p_1 + p_2$  we get

$$\begin{aligned} (f + g)^*(p) &= \sup_x [\langle x, p \rangle - f(x) - g(x)] \\ &= \sup_x [\langle x, p_1 \rangle + \langle x, p_2 \rangle - f(x) - g(x)] \\ &\leq \sup_x [\langle x, p_1 \rangle - f(x)] + \sup_x [\langle x, p_2 \rangle - g(x)] \\ &= f^*(p_1) + g^*(p_2). \end{aligned}$$

Taking the infimum over all such decomposition on the right hand side shows

$$(f + g)^* \leq f^* \square g^*.$$

Step 1: Now we claim that if the more restrictive condition

$$\bigcup_{\lambda \geq 0} \lambda(\text{dom}(f) - \text{dom}(g)) = \mathbb{R}^d \quad (Q')$$

holds, then  $f^* \square g^*$  is lsc on  $\mathbb{R}^d$ . Let  $\mu \in \mathbb{R}$  and

$$C := \text{lev}_\mu(f^* \square g^*) = \{p \in \mathbb{R}^d \mid (f^* \square g^*)(p) \leq \mu\}$$

and aim to show that  $C$  is closed. For  $\epsilon > 0$  consider

$$C_\epsilon := \{q + r \in \mathbb{R}^d \mid f^*(q) + g^*(r) \leq \mu + \epsilon\}.$$

By definition of the infimal convolution we have  $(f^* \square g^*)(p) \leq \mu$  if and only if  $p = q + r$  such that for all  $\epsilon > 0$  it holds that  $f^*(q) + g^*(r) \leq \mu + \epsilon$ . In other words: It holds that

$$C = \bigcap_{\epsilon > 0} C_\epsilon$$

and hence, it is enough, to prove that all the  $C_\epsilon$  are closed. To that end, we consider the sets

$$\begin{aligned} K &= C_\epsilon \cap \overline{B_t(0)} \\ &= \{q + r \in \mathbb{R}^d \mid f^*(q) + g^*(r) \leq \mu + \epsilon, \|q + r\| \leq t\}. \end{aligned}$$

If all these  $K$  are closed, then  $C_\epsilon$  is closed. Let

$$H = \{(q, r) \in \mathbb{R}^d \times \mathbb{R}^d \mid f^*(q) + g^*(r) \leq \mu + \epsilon, \|q + r\| \leq t\}.$$

Since the map  $(q, r) \mapsto f^*(q) + g^*(r)$  is closed (both  $f^*$  and  $g^*$  are lsc), we see that  $H$  is closed.

We show that  $H$  is bounded: To show this, we show that there is a constant  $C(x, y)$  such that for all  $(q, r) \in H$  it holds that  $\langle x, q \rangle + \langle y, r \rangle \leq C(x, y)$ . By assumption (Q') we can write every  $(x, y)$  as

$$x - y = \lambda(u - v)$$

with some  $u \in \text{dom}(f)$ ,  $v \in \text{dom}(g)$  and  $\lambda \geq 0$ . Then (using the inequality by Fenchel and Cauchy-Schwarz)

$$\begin{aligned} \langle x, q \rangle + \langle y, r \rangle &= \lambda \langle u, q \rangle + \lambda \langle v, r \rangle + \langle y - \lambda v, q + r \rangle \\ &\leq \lambda(f^*(q) + f(u) + g^*(r) + g(v)) + \|q + r\| \|y - \lambda v\| \\ &\leq \lambda(\mu + \epsilon + f(u) + g(v)) + t \|y - \lambda v\| = C(x, y). \end{aligned}$$

This shows that  $H$  is bounded, and hence, compact. It remains to note that  $K$  is the image of  $H$  under the linear map  $(q, r) \mapsto q + r$  and hence,  $K$  is also compact, hence closed.

Step 2: Now we prove that if condition (Q') is fulfilled, we have  $(f + g)^* = f^* \square g^*$  and that the inf-convolution is exact. By Lemma 14.3 we have  $(f^* \square g^*)^* = f^{**} + g^{**} = f + g$  and another conjugation shows

$$(f + g)^* = (f^* \square g^*)^{**},$$

But our previous step showed that  $f^* \square g^*$  is lsc and hence  $(f^* \square g^*)^{**} = f^* \square g^*$ .

To see that the inf-convolution is exact, we note that we can see (similar to the first step) that for each  $\mu$  the set  $\{q \mid f^*(q) + g^*(p - q) \leq \mu\}$  is closed and hence, the infimum in the definition of the inf-convolution is attained.

Step 3: In the last step we get rid of the restrictive assumption (Q'). We use the following fact: If  $A \subset \mathbb{R}^d$  is convex and  $\bigcup_{\lambda \geq 0} \lambda A$  is a subspace, then  $0 \in A$  and hence  $\bigcup_{\lambda > 0} \lambda A = \bigcup_{\lambda \geq 0} \lambda A$ .

Let  $a \in A$ . Since  $a \in \bigcup_{\lambda \geq 0} \lambda A$  and the latter set is a vector space, there is  $b \in A$  such that  $-a = \lambda b$ . But then we have that

$$\frac{1}{1+\lambda}a + \frac{\lambda}{1+\lambda}b \in A$$

but the convex combination on the right hand side is 0.

From Assumption (Q) and the previous observation it follows that  $\text{dom}(f) \cap \text{dom}(g) \neq \emptyset$  and we may assume without loss of generality that  $0 \in \text{dom}(f) \cap \text{dom}(g)$ . We define the subspace  $V = \bigcup_{\lambda \geq 0} \lambda(\text{dom}(f) - \text{dom}(g))$  and note  $\text{dom}(f) \subset V$  and  $\text{dom}(g) \subset V$ . Hence, we could have worked in  $V$  from the start and since  $V$  is isomorphic to  $\mathbb{R}^n$ , the proof is complete.  $\square$

As a consequence, the subgradient sum-rule also holds if (Q) is fulfilled:

**Corollary 14.6.** Let  $f$  and  $g$  be proper, convex and lsc and fulfill condition (Q). Then it holds that

$$\partial(f + g) = \partial f + \partial g.$$

As we have seen, the inclusion  $\partial f + \partial g \subset \partial(f + g)$  always holds. For the reverse inclusion let  $p \in \partial(f + g)(x)$ . By Fenchel's equality (Lemma 13.3) we get

$$(f + g)(x) + (f + g)^*(p) = \langle p, x \rangle.$$

Theorem 14.5 shows  $(f + g)^* = f^* \square g^*$  and that the inf-convolution is exact, i.e. we have  $(f + g)^*(p) = f^*(p - q) + g^*(q)$  for some  $q$ . We get

$$f(x) + g(x) + f^*(p - q) + g^*(q) = \langle p - q, x \rangle + \langle q, x \rangle.$$

We conclude  $p - q \in \partial f(x)$  and  $q \in \partial g(x)$  (since  $q \notin \partial g(x)$  would imply  $g(x) + g^*(q) > \langle x, q \rangle$  from which we would get  $f(x) + f^*(p - q) < \langle p - q, x \rangle$  which contradicts Fenchel's inequality). Hence, we have  $p \in \partial f(x) + \partial g(x)$ .

Finally, let us note that condition (Q) is more general than the assumption in Theorem 10.4 namely that

$$\exists x \in \text{dom}(f) \cap \text{dom}(g), f \text{ continuous at } x \implies \text{(Q) fulfilled.}$$

If there is  $x \in \text{dom}(f) \cap \text{dom}(g)$  and  $f$  is continuous at  $x$ , then  $x \in \text{int dom}(f)$ , i.e.  $B_\epsilon(x) \subset \text{dom}(f)$  for some  $\epsilon$ . Hence  $\text{dom}(f) - \text{dom}(g) \supset B_\epsilon(x) - \{x\} = B_\epsilon(0)$  and we see that  $\bigcup_{\lambda \geq 0} \lambda(\text{dom}(f) - \text{dom}(g)) = \mathbb{R}^d$  while for (Q) we only need that the union is a subspace.



## 15 Fenchel-Rockafellar duality

The Fenchel equality states that

$$p \in \partial f(x) \iff f(x) + f^*(p) = \langle p, x \rangle \iff x \in \partial f^*(p).$$

Hence, we see that if  $x^*$  is a minimizer of  $f$ , then we know that  $x \in \partial f^*(0)$ . In other words: The subgradient of the conjugate at zero shows us, where minimizers of  $f$  are. Hence, knowing conjugate functions is quite helpful to treat minimization problems.

In this section, we use conjugate functions to derive a quite general notion of duality between optimization problem (which includes the notion of duality of linear programs, for example).

The problems which we will treat in this section are of the form

$$\min_{x \in \mathbb{R}^n} f(x) + g(Ax)$$

where  $A \in \mathbb{R}^{m \times n}$  and  $f : \mathbb{R}^n \rightarrow \bar{\mathbb{R}}$  and  $g : \mathbb{R}^m \rightarrow \bar{\mathbb{R}}$  are two proper, convex and lsc functions. We have seen examples for this, e.g. in Example 12.5 where we minimized  $\frac{1}{2}\|Ax - b\|_2^2 + \alpha R(x)$ , i.e. we could take  $f = R$  and  $g(y) = \frac{1}{2}\|y - b\|_2^2$ .

To motivate the duality, we express  $g$  via its conjugate and get

$$\begin{aligned} \min_{x \in \mathbb{R}^n} f(x) + g(Ax) &= \min_{x \in \mathbb{R}^n} f(x) + \sup_{y \in \mathbb{R}^m} \langle y, Ax \rangle - g^*(y) \\ &= \min_{x \in \mathbb{R}^n} \sup_{y \in \mathbb{R}^m} f(x) + \langle y, Ax \rangle - g^*(y). \end{aligned}$$

If we would know that the supremum was a maximum and that we could swap minimum and maximum, this would be equal to

$$\begin{aligned} \max_{y \in \mathbb{R}^m} \min_{x \in \mathbb{R}^n} f(x) + \langle y, Ax \rangle - g^*(y) &= \max_{y \in \mathbb{R}^m} \min_{x \in \mathbb{R}^n} [f(x) + \langle A^T y, x \rangle] - g^*(y) \\ &= \max_{y \in \mathbb{R}^m} - [\max_{x \in \mathbb{R}^n} \langle -A^T y, x \rangle - f(x)] - g^*(y) \\ &= \max_{y \in \mathbb{R}^m} -f^*(-A^T y) - g^*(y). \end{aligned}$$

The problem in the last line is called the (*Fenchel-Rockafellar*) *dual problem*. Note that the problem is a concave maximization problem, but since it is equivalent to

$$\min_{y \in \mathbb{R}^m} f^*(-A^T y) + g^*(y)$$

it is of exactly the same type than the problem we started with (and this problem is called *primal problem* in this context).

Let us explore the relation of the primal and dual problem in general. We start from middle ground, namely with a so-called *saddle point problem*, i.e. we have a function  $L : \mathbb{R}^n \times \mathbb{R}^m \rightarrow \bar{\mathbb{R}}$  and

want to find a pair  $(x^*, y^*)$  such that

$$\begin{aligned} x^* &\in \operatorname{argmin}_{x \in \mathbb{R}^n} L(x, y^*) \\ y^* &\in \operatorname{argmax}_{y \in \mathbb{R}^m} L(x^*, y). \end{aligned}$$

Any such pair will be called *saddle point* of  $L$ . Put differently, the solution  $(x^*, y^*)$  should satisfy

$$\forall x, y : L(x^*, y) \leq L(x, y) \leq L(x, y^*). \quad (*)$$

**Proposition 15.1.** For any  $L$  it holds the min-max inequality

$$\inf_x \sup_y L(x, y) \geq \sup_y \inf_x L(x, y).$$

Moreover,  $(x^*, y^*)$  is a saddle point of  $L$  if and only if

$$\min_x \max_y L(x, y) = \max_y \min_x L(x, y) = L(x^*, y^*).$$

*Proof.* For all  $\bar{x}, \bar{y}$  it holds that

$$\sup_y L(\bar{x}, y) \geq L(\bar{x}, \bar{y}) \geq \inf_x L(x, \bar{y}).$$

Taking the infimum over all  $\bar{x}$  on the left and the supremum over all  $\bar{y}$  on the right shows the inequality.

Now, let  $(x^*, y^*)$  be a saddle point. From the formulation (\*) above and the min-max inequality we get that

$$\begin{aligned} L(x^*, y^*) &\geq \max_y L(x^*, y) \geq \inf_x \max_y L(x, y) \\ &\geq \sup_y \inf_x L(x, y) = \sup_y \min_x L(x, y) \geq \min_x L(x, y^*) = L(x^*, y^*). \end{aligned}$$

Conversely, assume that the interchange of minimum and maximum gives the same result. Then

$$\min_x L(x, y^*) \leq L(x^*, y^*) \leq \max_y L(x^*, y)$$

which shows that  $(x^*, y^*)$  is a saddle point.  $\square$

*Example 15.2.* In many cases, one only has the min-max inequality, but no saddle points exist, even though maxima and minima exist. The simplest example may be

$$L(x, y) = \sin(x + y).$$

It holds that

$$\inf_y \sup_x \sin(x + y) = 1 > -1 = \sup_x \inf_y \sin(x + y)$$

even though the infima and suprema are attained.  $\triangle$

The statement should be read as “ $(x^*, y^*)$  is a saddle point, if and only if the minima and maxima exist and the equality holds”.

Check, that the change from sup to max and inf to min is justified in all places.

**Definition 15.3.** For a saddle-point problem with function  $L$  define  $F(x) = \sup_y L(x, y)$  and  $G(y) = \inf_x L(x, y)$ . Then the corresponding *primal* and *dual problem* are

$$\min_x F(x) \quad \text{and} \quad \max_y G(y),$$

respectively.

One sees that if  $(x^*, y^*)$  is a saddle point of  $L$ , then  $x^*$  solves the primal problem and  $y^*$  solves the dual problem and one has that  $F(x^*) = G(y^*)$ , i.e. the primal and dual optimal values coincide.

A little more terminology: If we have

$$\inf_x \sup_y L(x, y) = \sup_y \inf_x L(x, y)$$

we say that *strong duality* holds for the saddle point problem, while the min-max inequality shows that one always has *weak duality*. In terms of primal and dual problems: The primal and dual problems of a saddle point problem always obey *weak duality* i.e. it always holds that  $\inf_x F(x) \geq \sup_y G(y)$  and if  $\inf_x F(x) = \sup_y G(y)$ , even *strong duality* holds. Note that strong duality does not imply that the infimum or supremum are attained.

If a saddle point problem does not obey strong duality, we say that there is a *duality gap* and the difference  $\inf_x \sup_y L(x, y) - \sup_y \inf_x L(x, y)$  is called value of the *duality gap*.

Coming back to the problem

$$\min_x f(x) + g(Ax)$$

from the beginning of the section we see that this is the primal problem of the saddle point problem for

$$L(x, y) = f(x) + \langle Ax, y \rangle - g^*(y)$$

and the respective dual problem is

$$\max_y -f^*(-A^T y) - g^*(y).$$

In this context, the function  $L$  is also called *Lagrangian* of the problem. Weak duality

$$\inf_x f(x) + g(Ax) \geq \sup_y -f^*(-A^T y) - g^*(y)$$

always holds. Moreover, we will denote the *primal objective* by  $F(x) = f(x) + g(Ax)$  and the *dual objective* by  $G(y) = -f^*(-A^T y) - g^*(y)$ .

The knowledge of the dual problem is useful, to get cheap estimates on the distance to optimality:

**Proposition 15.4.** For convex, proper and lsc function  $f$  and  $g$  and matrix  $A$  define the gap function

$$\text{gap}(x, y) := f(x) + g(Ax) + f^*(-A^T y) + g^*(y) = F(x) - G(y).$$

Moreover, denote the primal and dual objective by  $F$  and  $G$  as above. Then it holds for any pair  $(\bar{x}, \bar{y})$  where  $G(\bar{y}) > -\infty$  that

$$\text{gap}(\bar{x}, \bar{y}) \geq F(\bar{x}) - \inf_x F(x).$$

*Proof.* By weak duality one has  $F(\bar{x}) \geq \inf_x F(x) \geq \sup_y G(y) \geq G(\bar{y})$  and especially  $\inf_x F(x) \geq G(\bar{y})$ . Hence, we have

$$F(\bar{x}) - \inf_x F(x) \leq F(\bar{x}) - G(\bar{y}) = \text{gap}(\bar{x}, \bar{y}).$$

□

What is the corresponding statement for the distance to dual optimality?

**Theorem 15.5** (Fenchel-Rockafellar duality). Let  $f : \mathbb{R}^n \rightarrow \bar{\mathbb{R}}, g : \mathbb{R}^m \rightarrow \bar{\mathbb{R}}$  be proper, convex and lsc and let  $A \in \mathbb{R}^{n \times m}$ . If

$$\bigcup_{\lambda \geq 0} \lambda(\text{dom}(g) - A \text{dom}(f)) = \mathbb{R}^m$$

then strong duality holds, i.e.

$$\inf_{x \in \mathbb{R}^n} f(x) + g(Ax) = \max_{y \in \mathbb{R}^m} -f^*(-A^T y) - g^*(y)$$

(and especially the max on the right hand side is attained).

*Proof.* We define  $\Phi : \mathbb{R}^n \times \mathbb{R}^m \rightarrow \bar{\mathbb{R}}$  by  $\Phi(x, y) = f(x) + g(y)$  and let  $M = \{(x, Ax) \mid x \in \mathbb{R}^n\}$  (i.e.  $M$  is the graph of  $A$ ).

We aim to show that

$$\bigcup_{\lambda \geq 0} \lambda(\text{dom}(\Phi) - M) = \mathbb{R}^n \times \mathbb{R}^m.$$

To that end, let  $(x, y) \in \mathbb{R}^n \times \mathbb{R}^m$ . By assumption, there exists  $\lambda \geq 0, u \in \text{dom}(f)$  and  $v \in \text{dom}(g)$  such that

$$y - Tx = \lambda(v - Tu)$$

and one can show (as in the beginning of Step 3 in Theorem 14.5) that one can even choose  $\lambda > 0$ . If we set  $a = u - x/\lambda$  we get

$$x = \lambda(u - a), \quad y = \lambda(v - Ta)$$

and this shows that indeed  $(x, y) \in \bigcup_{\lambda \geq 0} \lambda(\text{dom}(\Phi) - M)$ .

Now we define  $\Psi(x, y) = I_M(x, y)$  and note that we have just shown that condition (Q) is fulfilled for  $\Phi$  and  $\Psi$ . Thus we have by Theorem 14.5

$$(\Phi + \Psi)^* = \Phi^* \square \Psi^*$$

and the infimal convolution is exact. Actually, we only need this equality at 0 since

$$\begin{aligned} (\Phi + \Psi)^*(0) &= \sup_{(x,y)} -\Phi(x,y) - \Psi(x,y) = \sup_{y=Ax} -f(x) - g(y) \\ &= -\inf_x f(x) + g(Ax) \end{aligned}$$

and

$$\begin{aligned} (\Phi^* \square \Psi^*)(0) &= \min_{(x,y)} \Phi^*(x,y) + \Psi^*(-(x,y)) = \min_{x=-A^T y} f^*(x) + g^*(y) \\ &= \min_y f^*(-A^T y) + g^*(y). \end{aligned}$$

□

## 16 Examples of duality and optimality systems

*Example 16.1 (LP duality).* A *linear program* is an optimization problem where the objective function is linear and where there are linear equality and inequality constraints. The standard form of a linear program is: Given  $c \in \mathbb{R}^n$ ,  $A \in \mathbb{R}^{m \times n}$  and  $b \in \mathbb{R}^m$  solve

$$\min_{x \in \mathbb{R}^n} \langle c, x \rangle \quad \text{subject to} \quad Ax \leq b$$

where the inequality is understood componentwise. In other words

$$\min_{x \in \mathbb{R}^n} \{ \langle c, x \rangle \mid Ax \leq b \}.$$

We rewrite this in the context of Fenchel-Rockafellar duality as

$$\min_{x \in \mathbb{R}^n} f(x) + g(Ax)$$

with  $f(x) = \langle c, x \rangle$ ,  $g(v) = I_{\mathbb{R}_{\leq 0}^m}(v - b)$ . The conjugates are

$$f^*(p) = \sup_x \langle p - c, x \rangle = \begin{cases} 0, & \text{if } p - c = 0 \\ \infty, & \text{else,} \end{cases} = I_{\{c\}}(p),$$

$$g^*(y) = \sup_{v-b \leq 0} \langle v, y \rangle = \begin{cases} \langle b, y \rangle, & \text{if } y \geq 0 \\ \infty, & \text{else.} \end{cases}$$

Hence, the dual problem is

$$\max_{y \in \mathbb{R}^m} -f^*(-A^T y) - g^*(y) = \max_y \{ -\langle b, y \rangle \mid -A^T y = c, y \geq 0 \}$$

I.e. the (Fenchel-Rockafellar) dual of a linear problem is another linear program, namely

$$\max_{y \in \mathbb{R}^m} -\langle b, y \rangle \quad \text{subject to} \quad \begin{aligned} A^T y + c &= 0 \\ y &\geq 0. \end{aligned}$$

If  $m < n$ , then the dual problem has fewer variables (but more constraints).  $\triangle$

*Example 16.2 (Equality constrained norm minimization).* We consider the primal problem

$$\min_{x \in \mathbb{R}^n} \|x\| \quad \text{subject to} \quad Ax = b$$

with  $A \in \mathbb{R}^{m \times n}$ ,  $b \in \mathbb{R}^m$  and  $\|\cdot\|$  denoting any norm on  $\mathbb{R}^n$ . With  $f(x) = \|x\|$  and  $g(v) = I_{\{b\}}(v)$  this is of the form  $\min_x f(x) + g(Ax)$ . The conjugate of  $g$  is simply

$$g^*(y) = \langle b, y \rangle$$

and the conjugate of  $f$  is

$$f^*(p) = \sup_x \langle p, x \rangle - \|x\|.$$

With the notion of *dual norm*, defined by

$$\|p\|_* = \sup_{\|x\| \leq 1} \langle p, x \rangle$$

which fulfills  $\langle p, x \rangle \leq \|x\| \|p\|_*$  we can express the conjugate of  $f$  as

$$f^*(p) = I_{\{\|\cdot\|_* \leq 1\}}(p).$$

Hence, the dual problem is

$$\max_{y \in \mathbb{R}^m} -\langle b, y \rangle \quad \text{subject to} \quad \|A^T y\|_* \leq 1.$$

In the case of the 1-norm (whose dual is the  $\infty$ -norm), the primal

$$\min_{Ax=b} \|x\|_1$$

has the dual

$$\max_{\|A^T y\|_\infty \leq 1} -\langle b, y \rangle$$

which can be written as a linear program

$$\begin{aligned} \max_{y \in \mathbb{R}^m} -\langle b, y \rangle \quad \text{subject to} \quad & A^T y \leq \mathbb{1} \\ & -A^T y \leq \mathbb{1}. \end{aligned}$$

By the previous example, we know that the primal should also be a linear program, can you see how to write it a such?

△

If both the subgradient sum-rule and the subgradient chain-rule hold for the objective

$$\min_x f(x) + g(Ax)$$

an optimal  $x^*$  is characterized by the inclusion

$$0 \in \partial f(x^*) + A^T \partial g(Ax^*).$$

Fenchel-Rockafellar duality allows for an alternative optimality system that uses the dual variable:

**Proposition 16.3.** *Let  $f, g$  be proper, convex and lower semicontinuous and assume strong duality is fulfilled and that the primal problem has a solution, i.e. we have*

$$\min_x f(x) + g(Ax) = \max_y -f^*(-A^T y) - g^*(y)$$

*then a pair  $(x^*, y^*)$  is a saddle point of  $f(x) + \langle Ax, y \rangle - g^*(y)$  if and only if*

$$\begin{aligned} -A^T y^* &\in \partial f(x^*) \\ y^* &\in \partial g(Ax^*). \end{aligned}$$

By the subgradient inversion theorem (Lemma 13.3), the primal dual optimality system is also equivalent to

$$\begin{aligned} -A^T y^* &\in \partial f(x^*) \\ Ax^* &\in \partial g^*(y^*). \end{aligned}$$

*This primal-dual optimality system (or Fenchel-Rockafellar duality system) is also equivalent to  $x^*$  being a solution to the primal problem and  $y^*$  being a solution to the dual problem.*

*Proof.* A pair  $(x^*, y^*)$  is optimal if and only if

$$-f^*(-A^T y^*) - g^*(y^*) = f(x^*) + g(Ax^*)$$

to which we add and subtract  $\langle y^*, Ax^* \rangle$  to get

$$\langle -A^T y^*, x^* \rangle + \langle y^*, Ax^* \rangle - f^*(-A^T y^*) - g^*(y^*) = f(x^*) + g(Ax^*).$$

By Fenchel's inequality we have  $\langle -A^T y^*, x^* \rangle \leq f(x^*) + f^*(-A^T y^*)$  and  $\langle y^*, Ax^* \rangle \leq g(Ax^*) + g^*(y^*)$ , and see that the previous equality is equivalent to

$$\langle -A^T y^*, x^* \rangle = f(x^*) + f^*(-A^T y^*) \quad \text{and} \quad \langle y^*, Ax^* \rangle = g(Ax^*) + g^*(y^*)$$

which, by Fenchel's equality, is equivalent to the primal-dual optimality system.  $\square$

*Example 16.4* (Primal-dual optimality for LPs). Let us work out the primal-dual optimality system for the LP from Example 16.1. The subgradient of  $f$  is just  $\partial f(x) = c$  (independent of  $x$ ). The subgradient of  $g$  fulfills

$$\partial g(v) = \begin{cases} 0, & \text{if } v < b, \\ \emptyset, & \text{if } v \not\leq b \end{cases}$$

In the remaining (interesting) cases where some of the inequalities  $v_i \leq b_i$  are tight, we have for any  $w \in \partial g(v)$  that  $w_i \in ]-\infty, 0]$ . Hence, the primal-dual optimality system is

$$\begin{aligned} -A^T y^* \in \partial f(x^*) &\implies -A^T y^* = c \\ y^* \in \partial g(Ax^*) &\implies Ax^* \leq b, \text{ and } \begin{cases} y_i^* = 0, & \text{if } (Ax^*)_i < b_i \\ y_i^* \leq 0, & \text{if } (Ax^*)_i = b_i. \end{cases} \end{aligned}$$

The last condition is a so-called *complementarity condition* and it states that at least one of the quantities  $y_i$  or  $(Ax - b)_i$  has to be zero for every  $i$ .

$\triangle$

*Example 16.5* (Primal-dual optimality system for constrained norm minimization). In the norm minimization example (Example 16.2), we have  $\partial g^*(v)$  is empty if  $v \neq b$ , but equal to  $\mathbb{R}$  for  $v = b$ . In total we get as optimality system

$$\begin{aligned} -A^T y^* \in \partial f(x^*) &\implies -A^T y^* = \partial \|x^*\| \\ y^* \in \partial g(Ax^*) &\implies Ax^* = b. \end{aligned}$$

Let's consider special cases:

- Let  $\|x\| = \|x\|_1$ . Then the subgradient fulfills

$$p \in \partial \|x\|_1 \iff \begin{cases} p_i = \text{sign}(x_i), & \text{if } x_i \neq 0 \\ |p_i| \leq 1, & \text{if } x_i = 0. \end{cases}$$

Hence, the primal-dual optimality system is

$$\begin{aligned} Ax^* &= b, \\ |A^T y|_i &\leq 1, \\ (A^T y)_i &= \text{sign}(x_i) \text{ if } x_i \neq 0. \end{aligned}$$



- In the case of the 2-norm one would rather take  $f(x) = \frac{1}{2}\|x\|_2^2$  (with subgradient  $\partial f(x) = \{x\}$ ) and get the primal-dual optimality system

$$\begin{aligned} -A^T y^* &= x^* \\ Ax^* &= b. \end{aligned}$$

△

In some cases the inclusion  $-A^T y \in \partial f(x)$  can help to recover a primal solution from a dual solution: By subgradient inversion (Lemma 13.3) the inclusion is equivalent to  $x \in \partial f^*(-A^T y)$ . If  $\partial f^*$  is single valued, this even leads to a single primal solution corresponding the any dual solution. The next propositions shows that this is the case, for example, when  $f$  is strongly convex.

**Proposition 16.6.** *Let  $f : \mathbb{R}^d \rightarrow \bar{\mathbb{R}}$  be proper, convex and lsc. Then the subgradient is monotone, in the sense that for  $p_i \in \partial f(x_i), i = 1, 2$ , then*

$$\langle p_1 - p_2, x_1 - x_2 \rangle \geq 0.$$

*If  $f$  is strongly convex with constant  $\mu$ , then it even holds that*

$$\langle p_1 - p_2, x_1 - x_2 \rangle \geq \mu \|x_1 - x_2\|_2^2.$$

*Proof.* The first claim follows by adding the two subgradient inequalities  $f(x_2) \geq f(x_1) + \langle p_1, x_2 - x_1 \rangle$  and  $f(x_1) \geq f(x_2) + \langle p_2, x_1 - x_2 \rangle$ . For the second claim, apply the first one to the convex function  $g(x) = f(x) - \frac{\mu}{2}\|x\|_2^2$  with  $p_i - \mu x_i \in \partial g(x_i)$ . □

**Proposition 16.7.** *If  $f : \mathbb{R}^d \rightarrow \bar{\mathbb{R}}$  is proper, strongly convex with constant  $\mu$  and lsc, then:*

- $\text{dom}(f^*) = \mathbb{R}^d$ ,
- $f^*$  is  $1/\mu$ -smooth and moreover  $\nabla f^*(p) = \text{argmax}_x \langle p, x \rangle - f(x)$ ,

Recall that  $f^*$  is  $L$ -smooth if it is differentiable with  $\nabla f^*$  being Lipschitz continuous with constant  $L$ .

*Proof.* i) Since  $f^*(p) = \sup_x \langle x, p \rangle - f(x)$ , we see that strong convexity of  $f$  ensures existence of maximizers for every  $p$ , and this gives a finite value for the supremum for every  $p$ .

- By Fermat's principle, some  $x$  maximizes  $\langle p, x \rangle - f(x)$  exactly if  $p \in \partial f(x)$  which is, by subgradient inversion, equivalent to  $x \in \partial f^*(p)$ . By strong convexity of  $f$ , the maximizer  $x$  is unique for every  $p$ , which shows that  $\partial f^*(p)$  is a singleton and Proposition 9.7 implies differentiability of  $f^*$  and

$$\partial f^*(p) = \{\nabla f^*(p)\} = \{x\}.$$

This also implies that  $\nabla f^*(p) = \text{argmax}_x \langle p, x \rangle - f(x)$ .

Proposition 16.6 shows for  $p \in \partial f(x)$  and  $p' \in \partial f(x')$  that

$$\langle p - p', x - x' \rangle \geq \mu \|x - x'\|^2.$$

Subgradient inversion gives  $x \in \nabla f^*(p)$  and  $x' \in \nabla f^*(p')$  and this gives

$$\langle p - p', \nabla f^*(p) - \nabla f^*(p') \rangle \geq \mu \|\nabla f^*(p) - \nabla f^*(p')\|_2^2.$$

Applying Cauchy-Schwarz's inequality shows  $\|\nabla f^*(p) - \nabla f^*(p')\|_2 \leq \frac{1}{\mu} \|p - p'\|_2$  as claimed.

□

## 17 Classes of optimization problems

We are going to develop a theory that will allow us to say how hard a certain class of optimization problems is. To that end we will have to pin down some ingredients:

- The problem class: How do we describe a problem from a class? What properties do we assume for a problem?
- The algorithm: What information of the problem can be used by the algorithm?
- A notion of approximate solution: How do we measure how good some approximate solution is?

We will analyze iterative algorithms and our goal is, to quantify how many steps are needed to get an answer with a given accuracy.

We will start with a very simple case (that is actually not related to convex analysis at all):

### Global optimization of Lipschitz functions on bounded domains.

**Problem class:** Let  $Q = [0, 1]^n$  be the unit cube and  $f : Q \rightarrow \mathbb{R}$  be Lipschitz continuous with constant  $L$  with respect to the  $\infty$ -norm. The problem we consider is

$$\min_{x \in Q} f(x).$$

**Oracle:** In each step we are able to query the function value  $f(x)$  for some point  $x \in Q$ .

Here is a very simple algorithm:

---

#### Algorithm 1: Grid search

---

**Input:**  $p \in \mathbb{N}$

For all  $(i_1, \dots, i_n) \in \{0, \dots, p\}^n$  assemble points

$$x_{(i_1, \dots, i_n)} = \left[ \frac{i_1}{p}, \dots, \frac{i_n}{p} \right]^T$$

In every point  $x_{(i_1, \dots, i_n)}$  evaluate the functional value;

Find  $\bar{x}$  among the  $x_{(i_1, \dots, i_n)}$ , which has the smallest objection value;

**Result:** Pair  $(\bar{x}, f(\bar{x}))$

---

Let us analyze this method:

**Theorem 17.1** (Upper complexity bound for grid search). *Denote  $f^* := \min_{x \in Q} f(x)$  and let  $\bar{x}$  be the output of Algorithm 1. Then it holds that*

$$f(\bar{x}) - f^* \leq \frac{L}{2^p}.$$

We consider Lipschitz continuous function for some reason: Since we want a certain accuracy for some given iterate  $x^k$ , we would like to be able to estimate  $f(x^k) - \min_{x \in Q} f(x)$  in some way. If  $f$  is merely continuous, the function may vary as much as it like in a small neighborhood, and with Lipschitz continuity we are able to bound such variations.

This is called a *zeroth order oracle* and methods that only use functions evaluations are called *zeroth order methods*.

*Proof.* Let  $f^* = f(x^*)$ . Then,  $x^*$  lies in one grid cell, i.e. we have

$$x := x_{(i_1, \dots, i_n)} \leq x^* \leq y_{(i_1+1, \dots, i_n+1)} =: y.$$

From these two surrounding points we define the best point on the grid nearby as

$$(\tilde{x})_i = \begin{cases} y_i, & x_i^* \geq \frac{x_i + y_i}{2} \\ x_i, & \text{else.} \end{cases}$$

Then we have  $|\tilde{x}_i - x_i^*| \leq 1/(2p)$  and since  $\tilde{x}$  is among the grid points, we have

$$f(\tilde{x}) - f^* \leq f(\tilde{x}) - f(x^*) \leq L \|\tilde{x} - x^*\|_\infty \leq \frac{L}{2p}. \quad \square$$

In other words: To reach an accuracy  $f(\tilde{x}) - f^* \leq \epsilon$  we need  $p \geq L/(2\epsilon)$ , and since the number of points in the grid is  $N = (p+1)^n$ , the number of function evaluations (or calls to the oracle) is at least

$$N \geq (\frac{L}{2\epsilon} + 1)^n.$$

While this is a lot, the next theorem shows, that one can not do much better than that. We can prove such result, by constructing a problem that tries to annoy a given problem as much as possible.

**Theorem 17.2.** *If  $\epsilon < L/2$ , there is no method that solves the problem with less than  $(\lfloor L/(2\epsilon) \rfloor)^n$  function evaluations.*

*Proof.* Set  $p = \lfloor L/(2\epsilon) \rfloor$  (note  $p \geq 1$ ) and assume that there is method that needs  $N < p^n$  function evaluations to solve any problem in our class. We construct an “annoying”  $f$ : We build  $f$  such that  $f(x) = 0$  for all points that have been tried, but  $f^* = \min f$  being as small as possible.

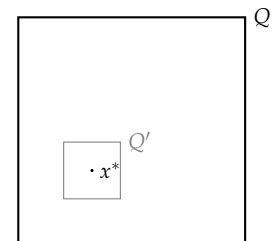
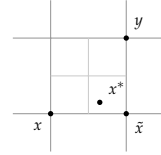
Since  $p \geq 1$ , and since we have  $N < p^n$  evaluation points, there has to be a small cube  $Q' \subset Q$  of sidelength  $1/p$  that does not contain any of these evaluation points. Let  $x^*$  be the midpoint of  $Q'$  and set

$$f(x) = \min(0, L\|x - x^*\|_\infty - \epsilon')$$

with  $\epsilon < \epsilon' < L/(2p)$ . This  $f$  is  $L$ -Lipschitz by definition and equals 0 outside of  $Q'$  (since  $\epsilon'/L < 1/(2p)$ ). Its minimal values is  $f^* = f(x^*) = -\epsilon'$ . Hence, the accuracy of this method is  $\epsilon' > \epsilon$ .  $\square$

We have seen that grid search need roughly  $(\frac{1}{\epsilon})^n$  iterations and the lower bound we got in the theorem is of the same order of magnitude. Hence, grid search is actually not a bad method for this class, although the number of iterations can be insanely large. This indicates, that general problems from this class can be very hard.

Here are further problems classes that we will deal with:



**Convex and  $L$ -Lipschitz.**

**Problem class:**  $f : \mathbb{R}^n \rightarrow \bar{\mathbb{R}}$ , proper, convex and  $L$ -Lipschitz with respect to the 2-norm and the problem is

$$\min_{x \in \mathbb{R}^n} f(x).$$

**Oracle:** At a given point  $x^k$  we are able to query one subgradient  $p^k \in \partial f(x^k)$ .

**Method:** The iterates will only move in the set

$$x^k \in x^0 + \text{span}\{p^0, \dots, p^{k-1}\}$$

i.e. in each step we can only move into directions of subgradients which we have already encountered.

**Convex and  $L$ -smooth.**

**Problem class:**  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  convex and differentiable with  $L$ -Lipschitz gradient and the problem is

$$\min_{x \in \mathbb{R}^n} f(x).$$

**Oracle:** At a given point  $x^k$  we are able to query the gradient  $g^k = \nabla f(x^k)$ .

**Method:** The iterates will only move in the set

$$x^k \in x^0 + \text{span}\{g^0, \dots, g^{k-1}\}$$

i.e. in each step we can only move into directions of gradients which we have already encountered.

**Convex,  $L$ -smooth and  $\mu$ -strongly convex.** Same as before, and we additionally assume that  $f$  is strongly convex with constant  $\mu > 0$ .

Now we collected a few more facts about smooth and strongly convex functions:

**Theorem 17.3.** *The following conditions are equivalent to  $f$  being convex and  $L$ -smooth (each holding for all  $x, y$  and  $\lambda \in [0, 1]$ ):*

- i)  $0 \leq f(y) - f(x) - \langle \nabla f(x), y - x \rangle \leq \frac{L}{2} \|x - y\|^2$
- ii)  $f(x) + \langle \nabla f(x), y - x \rangle + \frac{1}{2L} \|\nabla f(x) - \nabla f(y)\|^2 \leq f(y)$
- iii)  $\frac{1}{L} \|\nabla f(x) - \nabla f(y)\|^2 \leq \langle \nabla f(x) - \nabla f(y), x - y \rangle$
- iv)  $\langle \nabla f(x) - \nabla f(y), x - y \rangle \leq L \|x - y\|^2$
- v)  $\lambda f(x) + (1 - \lambda)f(y) \geq f(\lambda x + (1 - \lambda)y) + \frac{\lambda(1 - \lambda)}{2L} \|\nabla f(x) - \nabla f(y)\|^2$
- vi)  $\lambda f(x) + (1 - \lambda)f(y) \leq f(\lambda x + (1 - \lambda)y) + \frac{\lambda(1 - \lambda)L}{2} \|x - y\|^2$

The property in iii) is called *cocoercivity* of the gradient  $\nabla f$ . It is not to be confused with strong convexity where one has a lower bound of the form  $\mu/2 \|x - y\|^2$ .

*Proof.* Claim i) is just convexity combined with Lemma 12.4.

Now fix  $x_0$  and consider  $\varphi(y) = f(y) - \langle \nabla f(x_0), y \rangle$  which has its minimum at  $y^* = x_0$ . By i) we have

$$\begin{aligned}\varphi(y^*) &\leq \varphi(y - \tfrac{1}{L} \nabla \varphi(y)) \leq \varphi(y) + \tfrac{1}{2} \|\tfrac{1}{L} \nabla \varphi(y)\|^2 + \langle \nabla \varphi(y), -\tfrac{1}{L} \nabla \varphi(y) \rangle \\ &= \varphi(y) - \tfrac{1}{2L} \|\nabla \varphi(y)\|^2\end{aligned}$$

which shows ii) since  $\nabla \varphi(y) = \nabla f(y) - \nabla f(x_0)$ .

Inequality iii) follows from ii) by adding two copies of ii) with  $x$  and  $y$  swapped and by Cauchy-Schwarz we get  $L$ -smoothness from iii).

Inequality iv) follows from i) by adding two copies and vice versa iv) implies i) by

$$\begin{aligned}f(y) - f(x) - \langle \nabla f(x), y - x \rangle &= \int_0^1 \langle \nabla f(x + \tau(y - x)) - \nabla f(x), y - x \rangle d\tau \\ &\leq \tfrac{L}{2} \|y - x\|^2.\end{aligned}$$

To get v) from ii) set  $x_\alpha = \alpha x + (1 - \alpha)y$  and note

$$\begin{aligned}f(x) &\geq f(x_\alpha) + \langle \nabla f(x_\alpha), (1 - \alpha)(x - y) \rangle + \tfrac{1}{2L} \|\nabla f(x) - \nabla f(x_\alpha)\|^2 \\ f(y) &\geq f(x_\alpha) - \langle \nabla f(x_\alpha), \alpha(y - x) \rangle + \tfrac{1}{2L} \|\nabla f(y) - \nabla f(x_\alpha)\|^2.\end{aligned}$$

Multiplying by  $\alpha$  and  $(1 - \alpha)$ , respectively, and adding we get v). Conversely, we get ii) from v) by dividing by  $1 - \alpha$  and  $\alpha \rightarrow 1$ .

Similarly, one shows the equivalence of i) and vi).  $\square$

**Lemma 17.4.** A differentiable function  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  is  $\mu$ -strongly convex if and only if for all  $x, y$  we have

$$f(y) \geq f(x) + \langle \nabla f(x), y - x \rangle + \tfrac{\mu}{2} \|y - x\|_2^2.$$

*Proof.* Just apply Theorem 5.1 iii) to the convex function  $g(x) = f(x) - \tfrac{\mu}{2} \|x\|_2^2$  with gradient  $\nabla g(x) = \nabla f(x) - \mu x$ .  $\square$

**Theorem 17.5.** Let  $f$  be  $\mu$ -strongly convex. Then it holds for all  $x, y$  that

$$\begin{aligned}f(y) &\leq f(x) + \langle \nabla f(x), y - x \rangle + \tfrac{1}{2\mu} \|\nabla f(x) - \nabla f(y)\|_2^2 \\ \langle \nabla f(x) - \nabla f(y), x - y \rangle &\leq \tfrac{1}{\mu} \|\nabla f(x) - \nabla f(y)\|_2^2.\end{aligned}$$

*Proof.* For the first inequality we fix  $x$  and tilt  $f$  by defining  $\varphi(y) = f(y) + \langle \nabla f(x), y \rangle$ . Note that  $\nabla \varphi(x) = 0$  and thus  $\varphi$  is minimal at  $x$ . Since  $\varphi$  is still  $\mu$ -strongly convex, we have  $\varphi(x) \geq \min_z \varphi(z) \geq \min_z [\varphi(y) + \langle \nabla \varphi(y), z - y \rangle + \tfrac{\mu}{2} \|z - y\|_2^2]$ . We calculate that the minimum on the right is attained at  $z = y - \tfrac{1}{\mu} \nabla \varphi(y)$  and has the value  $\varphi(y) - \tfrac{1}{2\mu} \|\nabla \varphi(y)\|_2^2$  which is exactly the first inequality.

For the second inequality just swap the roles of  $x$  and  $y$  in the first and add both of them.  $\square$

**Theorem 17.6.** Let  $f$  be  $\mu$ -strongly convex and  $L$ -smooth with  $L \geq \mu$ . Then it holds for all  $x, y$  that

$$\langle \nabla f(x) - \nabla f(y), x - y \rangle \geq \frac{\mu L}{\mu + L} \|x - y\|_2^2 + \frac{1}{\mu + L} \|\nabla f(x) - \nabla f(y)\|_2^2.$$

*Proof.* We define the function  $\varphi(x) = f(x) - \frac{\mu}{2} \|x\|_2^2$  which is still convex and since the gradient is  $\nabla \varphi(x) = \nabla f(x) - \mu x$ , is it  $L - \mu$ -smooth.

In the case  $L = \mu$  we see from Theorem 17.3, iv) and Proposition 16.6 that  $\langle \nabla f(x) - \nabla f(y), x - y \rangle = \mu \|x - y\|_2^2$  from which we conclude  $f(x) = \frac{\mu}{2} \|x\|_2^2$  and the theorem holds.

For  $\mu < L$  we get from Theorem 17.3 iii) that

$$\langle \nabla \varphi(x) - \nabla \varphi(y), x - y \rangle \geq \frac{1}{L - \mu} \|\nabla \varphi(x) - \nabla \varphi(y)\|_2^2.$$

The left hand side evaluates to

$$\langle \nabla f(x) - \nabla f(y), x - y \rangle - \mu \|x - y\|_2^2$$

while the right hand side is

$$\frac{1}{L - \mu} (\|\nabla f(x) - \nabla f(y)\|_2^2 - 2\mu \langle \nabla f(x) - \nabla f(y), x - y \rangle + \mu^2 \|x - y\|_2^2).$$

Plugging this in and cleaning up proves the result.  $\square$

## 18 Convergence rates and worst case analysis

Now we will apply the concept of “annoying problems” to more classes of optimization problems. We start with the class of convex and  $L$ -Lipschitz functions, i.e. we do not assume differentiability or strong convexity and can only use subgradients in each iteration. Here is the result on worst case analysis:

**Theorem 18.1.** *For every  $k \in \{0, \dots, n+1\}$  there exists some convex function  $f_k : \mathbb{R}^n \rightarrow \bar{\mathbb{R}}$  which is Lipschitz continuous with constant  $L$ , and has a minimum  $f_k^* = f_k(x^*)$  at some  $x^*$  with  $\|x^*\|_2 \leq R$  and an oracle that gives a subgradient  $p \in \partial f(x)$  such that any sequence  $x^k$  that fulfills  $x^k \in x^0 + \text{span}\{p^0, \dots, p^{k-1}\}$  fulfills*

$$f_k(x^k) - f_k^* \geq \frac{LR}{2(1+\sqrt{k+1})}.$$

*Proof.* For some constant  $\mu, \gamma > 0$  (to be defined later) we set

$$f_k(x) = \gamma \max_{1 \leq i \leq k} x_i + \frac{\mu}{2} \|x\|_2^2.$$

The subdifferential is

$$\partial f_k(x) = \gamma \text{conv}\{e_i \mid i \in I(x)\} + \mu x$$

where  $I(x) := \{j \in \{1, \dots, k\} \mid x_j = \max_{1 \leq i \leq k} x_i\}$ . This function is Lipschitz continuous on any ball  $B_\rho(0)$ , since by the subgradient inequality we have for all  $p_k(x) \in \partial f_k(x)$  that

$$f_k(y) - f_k(x) \leq \langle p_k(y), y - x \rangle \leq \|p_k(y)\|_2 \|x - y\|_2 \leq (\mu\rho + \gamma) \|x - y\|_2$$

and hence  $L = \mu\rho + \gamma$  is a Lipschitz constant.

Solving the inclusion  $0 \in \partial f_k(x)$  we see that a minimizer is at  $x^{k*}$  given by

$$(x^{k*})_i = \begin{cases} -\frac{\gamma}{\mu k}, & \text{if } 1 \leq i \leq k \\ 0, & \text{else.} \end{cases}$$

The norm of the minimizer and the optimal value are

$$R_k := \|x^{k*}\|_2 = \sqrt{k \left(\frac{\gamma}{\mu k}\right)^2} = \frac{\gamma}{\mu \sqrt{k}}, \quad f_k^* := f_k(x^{k*}) = -\frac{\gamma^2}{\mu k} + \frac{\mu}{2} R_k^2 = -\frac{\gamma^2}{2\mu k}.$$

Let us initialize the method with  $x^0 = 0$  and see what we can get. We aim to show that the  $j$ -th iterate ( $j \leq k$ )  $x^j$  has all entries with indices  $i = j+1, \dots, n$  equal to zero. In the first step, our oracle gives us the subgradient  $p_0 = \gamma e_1$  (others would be possible, but this is the worst choice) and hence,  $x^1$  has all entries  $x_i^1$  equal to zero with the only possible exception of  $i = 1$  and this proves the case  $j = 1$ .

For an induction assume that  $x^j$  fulfills the assumption. Our oracle gives the subgradient  $p^j = \mu x^j + \gamma e_{i^*}$  with  $i^* \leq j+1$  (note



that the first  $j$  entries may all be negative so that  $i^* = j + 1$  is possible), and this shows the claim.

Hence, in the first  $k - 1$  steps, the objective value fulfills

$$f_k(x^i) \geq \max_{1 \leq j \leq k} x_j^i \geq 0.$$

Now we choose our constants as

$$\gamma = \frac{\sqrt{k}L}{1+\sqrt{k}}, \quad \mu = \frac{L}{(1+\sqrt{k})R}$$

and observe that

$$f_k^* = -\frac{\gamma^2}{2\mu k} = -\frac{LR}{2(1+\sqrt{k})}$$

and  $\|x^0 - x^*\|_2 = \frac{\gamma}{\mu\sqrt{k}} = R$ . Finally, we compute that the function is indeed Lipschitz continuous on the ball  $B_R(0)$  with constant  $\mu R + \gamma = L$  as desired.  $\square$

Hence, we conclude that no algorithm that only uses subgradient steps is able to solve *all* optimization problems with Lipschitz-continuous convex objective with less than  $\mathcal{O}(1/\sqrt{k})$  operations. One says: The iterations complexity of convex optimization is  $\mathcal{O}(1/\sqrt{k})$ .

Now let us analyze the next class of problems: Convex and  $L$ -smooth objectives. At each iteration  $x^k$  we can query the gradient  $g^k = \nabla f(x^k)$  and the  $k$ -th iterate is assumed to be in the set

$$x^0 + \text{span}\{g^0, \dots, g^{k-1}\}.$$

Here is an annoying objective that is difficult for all such methods: For given  $L > 0$  and  $0 \leq k \leq n$  let  $f_k : \mathbb{R}^n \rightarrow \mathbb{R}$  be defined by

$$f_k(x) = \frac{L}{4} \left[ \frac{1}{2}((x_1)^2 + \sum_{i=1}^k (x_i - x_{i+1})^2 + (x_k)^2) - x_1 \right].$$

We rewrite this objective with the matrix

$$D_k = \left[ \begin{array}{ccc|cc} -1 & & & & \\ & 1 & -1 & & \\ & & \ddots & \ddots & \\ & & & 1 & -1 \\ & & & & 1 \\ \hline & & & 0_{n-k-1,k} & 0_{n-k-1,n-k} \end{array} \right].$$

as

$$f_k(x) = \frac{L}{4} \left( \frac{1}{2} \|D_k x\|_2^2 + \langle e_1, x \rangle \right)$$

We will call any method that fulfills this assumption a *first order method*.

and we have

$$\nabla f_k(x) = \frac{L}{4}(D_k^T D_k x - e_1) \quad \text{and} \quad \nabla^2 f_k(x) = \frac{L}{4} D_k^T D_k.$$

One sees

$$A_k = D_k^T D_k = \left[ \begin{array}{cccc|c} 2 & -1 & & & \\ -1 & 2 & \ddots & & \\ & \ddots & & -1 & 0_{k,n-k} \\ & & -1 & 2 & \\ \hline & & & & 0_{n-k,k} \end{array} \right]$$

and hence, we get

$$\begin{aligned} 0 \leq \frac{L}{4} \|D_k s\|^2 &= \langle \nabla^2 f_k(x) s, s \rangle \\ &= \frac{L}{4} \left( (s^{(1)})^2 + \sum_{i=1}^{k-1} (s^{(i)} - s^{(i+1)})^2 + (s^{(k)})^2 \right) \\ &\leq \frac{L}{4} \left( (s^{(1)})^2 + \sum_{i=1}^{k-1} 2((s^{(i)})^2 + (s^{(i+1)})^2) + (s^{(k)})^2 \right) \\ &\leq L \sum_{i=1}^n (s^{(i)})^2 = L \|s\|^2. \end{aligned}$$

This shows  $0 \preceq \nabla^2 f_k(x) \preceq LI$  and we have proven the  $f_k$  is convex and  $L$ -smooth.

**Proposition 18.2.** *The function  $f_k$  has a minimizer*

$$x_i^* = \begin{cases} 1 - \frac{i}{k+1}, & \text{if } i = 1, \dots, k \\ 0, & \text{if } i = k+1, \dots, n. \end{cases}$$

and optimal value and norm

$$f_k^* = f_k(x^*) = \frac{L}{8}(-1 + \frac{1}{k+1}) \quad \text{and} \quad \|x^*\|_2^2 \leq \frac{k+1}{3}, \text{ respectively.}$$

Note that the entries of the minimizer (and hence also  $f_k^*$  and its norm as well) depend on the value  $k$ . We will make use of this in the following.

*Proof.* The optimality condition is

$$0 = \nabla f_k(x) = \frac{L}{4}(A_k x - e_1)$$

which is

$$\left[ \begin{array}{cccc|c} 2 & -1 & & & \\ -1 & 2 & \ddots & & \\ & \ddots & & -1 & 0_{k,n-k} \\ & & -1 & 2 & \\ \hline & & & & 0_{n-k,k} \end{array} \right] \begin{bmatrix} x_1 \\ \vdots \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

We see that  $x_i = 0$  for  $i = k+1, \dots, n$ . The first equation is  $2x_1 - x_2 = 1$ , which gives

$$x_2 = 2x_1 - 1.$$

Plugging this is the second equation  $-x_1 + 2x_2 - x_3 = 0$  gives

$$x_3 = 3x_1 - 2.$$

Proceeding in this way, we get for  $i = 2, \dots, k$  from the  $i - 1$ th equation that

$$x_i = ix_1 - (i - 1). \quad (*)$$

The  $k$ -th equation  $-x_{k-1} + 2x_k = 0$  is then

$$0 = -(k-1)x_1 + (k-2) + 2(kx_1 - (k-1)) = (k+1)x_1 - k.$$

which shows  $x_1 = 1 - 1/(k+1)$ . Plugging this in  $(*)$  shows the formula for the minimizer.

For the minimal value we just plug in and get

$$\begin{aligned} f_k^* &= f_k(x^*) = \frac{L}{4} \left( \frac{1}{2} \|Dx^*\|^2 - \langle x^*, e_1 \rangle \right) \\ &= \frac{L}{4} \left( \frac{1}{2} \underbrace{\langle A_k x^*, x^* \rangle}_{e_1} - \langle x^*, e_1 \rangle \right) = -\frac{L}{8} \langle x^*, e_1 \rangle \\ &= \frac{L}{8} \left( -1 + \frac{1}{k+1} \right). \end{aligned}$$

Finally, we estimate

$$\begin{aligned} \|x^*\|^2 &= \sum_{i=1}^n (x_i^*)^2 = \sum_{i=1}^k \left( 1 - \frac{i}{k+1} \right)^2 = \sum_{i=1}^k \left( 1 - \frac{2i}{k+1} + \frac{i^2}{(k+1)^2} \right) \\ &= k - \frac{2}{k+1} \underbrace{\sum_{i=1}^k i}_{= \frac{k(k+1)}{2}} + \frac{1}{(k+1)^2} \underbrace{\sum_{i=1}^k i^2}_{\leq \frac{(k+1)^3}{3}} \leq \frac{k+1}{3}. \end{aligned} \quad (3)$$

□

Now let us analyze how methods according to our definition perform for this particular function.

**Lemma 18.3.** *Let  $1 \leq p \leq n$  and  $x_0 = 0$ . Then it holds for every sequence  $x^k$  with*

$$x^k \in L_k := \text{span}\{\nabla f_p(x_0), \dots, \nabla f_p(x_{k-1})\}$$

*and  $k \leq p$  that  $x^k = (*, \dots, *, 0, \dots, 0)$ , i.e. only the first  $k$  entries can be different from zero.*

*Proof.* Recall that  $\nabla f_p(x) = \frac{L}{4}(A_p x - e_1)$ , and hence we see that  $\nabla f_p(x^0) = -\frac{L}{4}e_1$ , and hence,  $x^1$  fulfills the claim.

No proceed inductively: if  $x^k$  only has the first  $k$  entries non-zero, then, since  $A_p$  is tridiagonal,  $g^{k+1} = \nabla f_p(x^k)$  has only the first  $k+1$  entries different zero and the same holds for  $x^{k+1}$ . □

**Corollary 18.4.** *For every sequence  $x^k$ ,  $k = 0, \dots, p$  with  $x^0 = 0$  and  $x^k \in L_k$  it holds that  $f_p(x^k) \geq f_k^*$ .*

We only observe that since  $x^k \in L_k$ , it only has the first  $k$  components non-zero and thus, that  $f_p(x^k) = f_k(x^k) \geq f_k^*$ .

We used the estimate  $(k+1)^3 = \sum_{i=0}^k [(i+1)^3 - i^3] = \sum_{i=0}^k [3i^2 + 3i + 1] \geq 3 \sum_{i=1}^k i^2$ . The exact sum would be  $\sum_{i=1}^k i^2 = \frac{k(k+1)(2k+1)}{6}$  (which we will use in the next section).