LAGRANGIAN DUALITY

MARKUS GRASMAIR

1. Primal and dual problems

Assume that we are given a constrained optimisation problem of the form

$$\min_x f(x) \qquad \text{subject to } \begin{cases} c_i(x) = 0, & i \in \mathcal{E}, \\ c_i(x) \geq 0, & i \in \mathcal{I}. \end{cases} \tag{1}$$

Then its Lagrangian is the mapping $\mathcal{L}\colon \mathbb{R}^n \times \mathbb{R}^{\mathcal{E}\cup\mathcal{I}}$ defined as

$$\mathcal{L}(x, \lambda) = f(x) - \sum_{i\in\mathcal{E}\cup\mathcal{I}} \lambda_i c_i(x).$$

Now define the function $p\colon \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$,

$$p(x) := \max_{\substack{\lambda\in\mathbb{R}^{\mathcal{E}\cup\mathcal{I}} \\ \lambda_i\geq 0,\ i\in\mathcal{I}}} \mathcal{L}(x, \lambda).$$

Then it is easy to see that

$$p(x) = \begin{cases} f(x) & \text{if } c_i(x) = 0,\ i \in \mathcal{E},\ \text{and } c_i(x) \geq 0,\ i \in \mathcal{I}, \\ +\infty & \text{else.} \end{cases}$$

Indeed, if $c_i(x) = 0$, $i \in \mathcal{E}$, and $c_i(x) \geq 0$, $i \in \mathcal{I}$, then

$$\mathcal{L}(x, \lambda) = f(x) - \sum_{i\in\mathcal{I}} \lambda_i c_i(x) \leq f(x)$$

for all $\lambda_i \geq 0$, $i \in \mathcal{I}$. Moreover, we obtain equality by choosing $\lambda = 0$. On the other hand, if any of the equality constraints is not satisfied, say $c_i(x) > 0$ for some $i \in \mathcal{E}$, then $\mathcal{L}(x, \lambda)$ can be made arbitrarily large by letting λ_i tend to $-\infty$. Similarly, if $c_i(x) < 0$ for any $i \in \mathcal{E} \cup \mathcal{I}$, then again $\mathcal{L}(x, \lambda)$ can be made arbitrarily large by letting λ_i tend to $+\infty$.

This shows that solving the constrained optimisation problem (1) is equivalent to solving the unconstrained problem

$$\min_{x\in\mathbb{R}^n} p(x),$$

or, explicitly, the *primal problem*

$$\min_{x\in\mathbb{R}^n} \max_{\substack{\lambda\in\mathbb{R}^{\mathcal{E}\cup\mathcal{I}} \\ \lambda_i\geq 0,\ i\in\mathcal{I}}} \mathcal{L}(x, \lambda). \tag{P}$$

Now one defines the *dual problem* by exchanging the order of the minimum and the maximum in (P):

Definition 1. The *Lagrangian dual* of (P) is the optimisation problem

$$\max_{\substack{\lambda\in\mathbb{R}^{\mathcal{E}\cup\mathcal{I}} \\ \lambda_i\geq 0,\ i\in\mathcal{I}}} \min_{x\in\mathbb{R}^n} \mathcal{L}(x, \lambda). \tag{D}$$

Date: April 2019.

More precisely, we first define a function $q\colon \mathbb{R}^{\mathcal{E}\cup\mathcal{I}} \to \mathbb{R}\cup\{-\infty\}$ by

$$q(\lambda) := \min_{x\in\mathbb{R}^n} \mathcal{L}(x,\lambda),$$

and then maximise q with respect to λ subject to the constraint that the Lagrange parameters for the inequality constraints are non-negative. Note that the minimum in the definition of q is taken over *all* $x \in \mathbb{R}^n$ irrespective of the constraints.

Remark 2. If one wants to be accurate, one should always read the minima and maxima in the definitions of the primal and the dual problem as infima and suprema, respectively, as it is not clear that these optimisation problems actually have solutions. Indeed, in the case of the primal formulation, the maximisation problem with respect to λ has a solution if and only if x is feasible.

Definition 3. By

$$d := \min_{x\in\mathbb{R}^n} p(x) \;-\; \max_{\substack{\lambda\in\mathbb{R}^{\mathcal{E}\cup\mathcal{I}}\\ \lambda_i\ge 0,\ i\in\mathcal{I}}} q(\lambda)$$

we denote the duality gap for the primal dual pair (P) and (D).

Lemma 4 (weak duality)**.** *Let d be the duality gap for (P) and (D). Then $d \ge 0$.*

Proof. This is an immediate consequence of the definition of the primal and the dual problem, and the inequality

$$\sup_{y\in Y}\inf_{x\in X} h(x,y) \le \inf_{x\in X}\sup_{y\in Y} h(x,y),$$

which holds whenever X are Y are non-empty sets and $h\colon X\times Y \to \mathbb{R}\cup\{\pm\infty\}$ is any function. $\square$

Example 5. Consider the optimisation problem

$$\min_x -\frac{1}{1+x^2} \qquad\qquad \text{s.t. } x^2 \ge 1.$$

The obvious solutions to this problem are the points $x = \pm 1$, where the value of the objective function is $-1/2$.

Now we compute the dual of this problem: The Lagrangian is

$$\mathcal{L}(x,\lambda) = -\frac{1}{1+x^2} - \lambda(x^2-1).$$

For $\lambda > 0$, the term $-\lambda x^2$ dominates the Lagrangian and we have

$$q(\lambda) = \inf_x \mathcal{L}(x,\lambda) = -\infty.$$

On the other hand, for $\lambda = 0$ we have

$$q(0) = \inf_x \mathcal{L}(x,0) = \inf_x -\frac{1}{1+x^2} = -1.$$

Finally, for $\lambda < 0$ we have

$$q(\lambda) = \inf_x \mathcal{L}(x,\lambda) = \inf_x\Big(-\frac{1}{1+x^2} - \lambda(x^2-1)\Big) = -1+\lambda,$$

as the infimum is always attained at $x = 0$. Thus

$$q(\lambda) = \begin{cases} -\infty & \text{if } \lambda > 0,\\ -1+\lambda & \text{if } \lambda \le 0.\end{cases}$$

Since the dual problem is a maximisation problem, the function value $-\infty$ for $\lambda > 0$ effectively serves as a constraint $\lambda \le 0$. In addition, we have the constraint $\lambda \ge 0$ from the fact that we have an inequality constraint. Thus we obtain the dual problem

$$\max_\lambda -1+\lambda \qquad\qquad \text{s.t. } \lambda = 0$$

with the (only possible) solution $\lambda = 0$ and an objective value of -1.

Consequently, we have a (non-zero) duality gap

$$d = \min_{x\in\mathbb{R}} p(x) \;\; - \max_{\lambda\geq 0} q(\lambda) = -\frac{1}{2} - (-1) = \frac{1}{2}.$$

2. Strong duality

The most important situation is that where the duality gap is equal to zero, as in this case the dual problem can be used for solving the original (*primal*) problem. In order to arrive at such results, we have to introduce the notion of saddle points. Note that the definition below is somehow different from the standard notion of saddle points used in basic calculus classes in that we are interested in *global* optimality properties with respect to the different variables.

Definition 6. The point $(x^*, \lambda^*) \in \mathbb{R}^d \times \mathbb{R}^{\mathcal{E}\cup\mathcal{I}}$ with $\lambda_i^* \geq 0$, $i \in \mathcal{I}$, is a *saddle point* of the Lagrangian, if (or a *primal–dual solution* of (P)), if

$$\mathcal{L}(x^*, \lambda) \leq \mathcal{L}(x^*, \lambda^*) \leq \mathcal{L}(x, \lambda^*)$$

for all $(x, \lambda) \in \mathbb{R}^d \times \mathbb{R}^{\mathcal{E}\cup\mathcal{I}}$ with $\lambda_i \geq 0$, $i \in \mathcal{I}$.

That is, a saddle point is a maximiser with respect to the (feasible) dual variables and a minimiser with respect to the primal variables.

Proposition 7. *Assume that (x^*, λ^*) is a saddle point of the Lagrangian and that $\mathcal{L}(x^*, \lambda^*) \in \mathbb{R}$. Then x^* is a solution of (P), λ^* is a solution of (D), and the complementarity conditions $\lambda_i^* c_i(x^*) = 0$, $i \in \mathcal{E}\cup\mathcal{I}$ hold. If the functions f and c_i, $i \in \mathcal{E}\cup\mathcal{I}$, are $\mathcal{C}^1$, then x^* is a KKT point with Lagrange multiplier λ^*.*

Proof. We note first that, for every λ with $\lambda_i \geq 0$, $i \in \mathcal{I}$, we have

$$q(\lambda) = \min_x \mathcal{L}(x, \lambda) \leq \mathcal{L}(x^*, \lambda) \leq \mathcal{L}(x^*, \lambda^*) = \min_x \mathcal{L}(x, \lambda^*) = q(\lambda^*).$$

Here the third and fourth relation are consequences of the assumption that (x^*, λ^*) is a saddle point. This shows that λ^* solves (D).

Similarly,

$$p(x) = \max_{\substack{\lambda\in\mathbb{R}^{\mathcal{E}\cup\mathcal{I}}\\ \lambda_i\geq 0,\; i\in\mathcal{I}}} \mathcal{L}(x, \lambda) \geq \mathcal{L}(x, \lambda^*) \geq \mathcal{L}(x^*, \lambda^*) = \max_{\substack{\lambda\in\mathbb{R}^{\mathcal{E}\cup\mathcal{I}}\\ \lambda_i\geq 0,\; i\in\mathcal{I}}} \mathcal{L}(x^*, \lambda) = p(x^*),$$

showing that x^* solves (P). In particular, since $\mathcal{L}(x^*, \lambda^*)$ is finite, this implies that x^* is a feasible point.

Now assume that the complementarity condition does not hold. Since x^* is feasible, this implies that there exists $i \in \mathcal{I}$ such that $c_i(x^*) > 0$ and $\lambda_i^* > 0$. In this case, however, replacing λ_i^* with $\hat{\lambda}_i := 0$ increases the value of the Lagrangian (without changing x^*). This is a contradiction to the assumption that (x^*, λ^*) is a saddle point (again, this uses the assumption that $\mathcal{L}(x^*, \lambda^*)$ is finite).

Finally, if the functions f and c_i, $i \in \mathcal{E}\cup\mathcal{I}$, are $\mathcal{C}^1$, then the fact that x^* minimises $\mathcal{L}(\cdot, \lambda^*)$ implies that the first order optimality condition

$$\nabla_x \mathcal{L}(x^*, \lambda^*) = 0$$

holds. As a consequence, all KKT conditions are satisfied. □

Remark 8. Note that the converse in general does not hold. That is, if (x^*, λ^*) is a KKT point, it is not necessarily a saddle point of the Lagrangian. Indeed, it is by no means guaranteed that the Lagrangian has any saddle points at all.

This can be seen in the problem discussed in Example 5: Here the points $x^* = \pm 1$ are KKT points (and global solutions) with Lagrange multipliers $\lambda^* = 1/4$.

However, the points $(x^*, \lambda^*) = (\pm 1, 1/4)$ are not saddle points of the Lagrangian in the sense of Definition 6, as

$$\mathcal{L}(0, 1/4) = -1 < -\frac{1}{2} = \mathcal{L}(\pm 1, 1/4).$$

Also, the Lagrange multiplier $\lambda^* = 1/4$ does not solve the dual problem, which is only finite for $\lambda = 0$.

Theorem 9. *Assume that x^* is a solution of (P), λ^* is a solution of (D), and that the duality gap is zero. Then (x^*, λ^*) is a saddle point of the Lagrangian.*

In particular, the complementarity conditions hold and x^* is a KKT point with Lagrange multiplier λ^* provided that the functions f and c_i are $\mathcal{C}^1$.

Proof. Since the duality gap is zero and x^* and λ^* solve the primal and dual problems, respectively, we have that

$$p(x^*) = q(\lambda^*)$$

Thus we have for every x that

$$\mathcal{L}(x, \lambda^*) \geq \min_{\hat{x}} \mathcal{L}(\hat{x}, \lambda^*) = q(\lambda^*) = p(x^*) = \max_{\substack{\lambda \in \mathbb{R}^{\mathcal{E} \cup \mathcal{I}} \\ \lambda_i \geq 0,\ i \in \mathcal{I}}} \mathcal{L}(x^*, \lambda) \geq \mathcal{L}(x^*, \lambda^*).$$

Similarly we have for every λ with $\lambda_i \geq 0$, $i \in \mathcal{I}$, that

$$\mathcal{L}(x^*, \lambda) \leq \max_{\substack{\hat{\lambda} \in \mathbb{R}^{\mathcal{E} \cup \mathcal{I}} \\ \hat{\lambda}_i \geq 0,\ i \in \mathcal{I}}} \mathcal{L}(x^*, \hat{\lambda}) = p(x^*) = q(\lambda^*) = \min_x \mathcal{L}(x, \lambda^*) \leq \mathcal{L}(x^*, \lambda^*).$$

This shows that (x^*, λ^*) is a saddle point of the Lagrangian.

The other assertions follow from Proposition 7. □

3. Duality and convex programming

In the following, we will discuss the application of duality theory to convex programmes, that is, convex optimisation problems with concave and linear inequality constraints, and linear equality constraints. More precisely, we assume that we are given constraints of the form

$$\begin{aligned} c_i(x) &\geq 0, \quad i \in \mathcal{I}, \\ Ax &\geq b, \\ Cx &= d, \end{aligned}$$

where the functions $c_i \colon \mathbb{R}^n \to \mathbb{R}$ are concave, $A \in \mathbb{R}^{m \times n}$, $C \in \mathbb{R}^{\ell \times n}$ are matrices, and $b \in \mathbb{R}^m$, $d \in \mathbb{R}^\ell$ are vectors. The inequalities $Ax \geq b$ are understood componentwise.

Definition 10. We say that *Slater's constraint qualification* is satisfied, if there exists $x \in \mathbb{R}^n$ with $Ax \geq b$, $Cx = d$, and $c_i(x) > 0$ for all $i \in \mathcal{I}$.

Remark 11. In the specific situation of only linear inequality and equality constraints, Slater's constraint qualification is equivalent to the feasibility of the constraints. In the case of additional non-linear (but concave) constraints, the condition is somehow stonger.

Theorem 12 (strong duality for convex programmes)**.** *Assume that f is convex and that Slater's constraint qualification holds. Assume morever that*

$$\inf_x p(x) > -\infty$$

(that is, the primal problem is bounded). Then the dual problem has a solution λ^* and the duality gap is zero.

If in addition the primal problem has a solution x^*, then (x^*, λ^*) is a saddle point of the Lagrangian. If moreover the functions f and c_i are $\mathcal{C}^1$, then x^* is a KKT point with Lagrange multiplier λ^*.

Proof. See [1, Thm. 11.15]. Note that the second part of the Theorem is an immediate consequence of Theorem 9, once it has been established that the duality gap is zero. □

Example 13. Consider the linear programme

$$(L) \qquad \min_x c^T x \qquad \text{s.t. } Ax \geq b.$$

The corresponding Lagrangian is

$$\mathcal{L}(x, \lambda) = c^T x - \lambda^T (Ax - b).$$

Thus the dual objective function is

$$q(\lambda) = \min_x \bigl(c^T x - \lambda^T (Ax - b)\bigr) = \lambda^T b + \min_x (c - A^T \lambda)^T x = \begin{cases} \lambda^T b & \text{if } A^T \lambda = c, \\ -\infty & \text{if } A^T \lambda \neq c. \end{cases}$$

Thus we can write the dual problem as

$$(L') \qquad \max_\lambda b^T \lambda \qquad \text{s.t. } \begin{cases} \lambda \geq 0, \\ A^T \lambda = c. \end{cases}$$

Note that we have again a linear programme, but the roles of the objective and constraint are reversed.

Now we will apply the results of Theorem 12 to this situation: To that end, we note first that the objective function is convex (since it is linear), and that we only have linear constraints. As a consequence, Slater's constraint qualification is satisfied if and only if the problem is *primal feasible*, that is, there exists a point $x \in \mathbb{R}^n$ satisfying the primal constraints $Ax \geq b$. Now assume that the problem is primal feasible and *bounded*, that is,

$$\inf_{Ax \geq b} c^T x > -\infty.$$

Then Theorem 12 is applicable and it follows that the dual problem (L') has a solution λ^*. In addition, it can be shown (see Remark 14 below) that in such a situation, the primal problem (L) admits a solution x^* as well.

Thus the primal-dual pair (x^*, λ^*) satisfies the KKT conditions, which in this case can be written as

$$\begin{aligned} A^T \lambda &= c, \\ Ax &\geq b, \\ \lambda &\geq 0, \\ \lambda^T (Ax - b) &= 0. \end{aligned} \tag{2}$$

Conversely, if (x^*, λ^*) solve the system (2), then x^* solves (L), λ^* solves (L'), and (since the duality gap is zero) $c^T x^* = b^T \lambda^*$.

In addition, if (x, λ) is any *primal-dual feasible* pair, that is, if $Ax \geq b$, $A^T \lambda = c$, and $\lambda \geq 0$, then $c^T x \geq b^T \lambda$. If, actually, $c^T x = b^T \lambda$, then (x, λ) is a primal-dual solution.

Remark 14. We consider again the linear programme

$$(L) \qquad \min_x c^T x \qquad \text{s.t. } Ax \geq b$$

with dual

$$(L') \qquad \max_\lambda b^T\lambda \qquad \text{s.t. } \begin{cases} \lambda \geq 0, \\ A^T\lambda = c, \end{cases}$$

from Example 13.

Since the dual is again a linear programme, we can try to compute its dual (the *double-dual* of (L)), and expect it to be a linear programme again. The Lagrangian of the dual programme is

$$\mathcal{L}'(\lambda; y, s) = b^T\lambda - y^T(A^T\lambda - c) - \lambda^T s,$$

and thus we obtain the double-dual problem (note that we have a maximisation problem, and thus the Lagrange parameters for the inequality constraints have to be non-positive!)

$$\min_{\substack{y,\, s \\ s \leq 0}} \max_\lambda \big(b^T\lambda - y^T(A^T\lambda - c) - s^T\lambda\big).$$

This can be rewritten as the linear programme

$$\min_{y,s} c^T y \qquad \text{s.t. } \begin{cases} s \leq 0, \\ Ay + s = b. \end{cases}$$

The Lagrange parameter s in this problem can now be interpreted as a slack variable, and we see that this double-dual is equivalent to the problem

$$\min_y c^T y \qquad \text{s.t. } Ay \geq b,$$

which is again the primal problem.

Thus we have shown that, apart from possible slack variables, the double-dual of a linear programme is again the primal programme. In particular, we can apply Theorem 12 to the *dual programme*, and conclude in particular that the primal problem has a solution provided that the dual programme is feasible and bounded. This, however, is guaranteed if the primal programme is feasible and bounded, since in this case the dual programme actually has a solution. In addition, if the primal problem is unbounded, then it follows from weak duality that the value of the dual problem is $-\infty$. This is only possible, if the dual problem is infeasible. Similarly, if the dual problem is unbounded (above), then weak duality implies that the value of the primal problem is $+\infty$, which implies that the primal problem is infeasible.

Thus we obtain the following results (cf. [2, Thm. 13.1]):

- If the primal (or the dual) problem is feasible and bounded, then there exists a primal-dual solution.
- If the primal problem is unbounded, then the dual problem is infeasible.
- If the dual problem is unbounded, then the primal problem is infeasible.

4. Dual methods for non-smooth optimisation

We now will apply these ideas to the solution of non-smooth (ℓ^1-based) optimisation problems of the form

$$(P_1) \qquad \min_x \frac{1}{2}\|x - z\|_2^2 + \alpha\|Ax\|_1,$$

where $z \in \mathbb{R}^n$ is some given, fixed vector, $A \in \mathbb{R}^{m\times n}$ is some matrix, and $\alpha > 0$. An equivalent reformulation, which allows us to apply Lagrangian duality, is the problem

$$(\hat{P}_1) \qquad \min_{x,y} \frac{1}{2}\|x - z\|_2^2 + \alpha\|y\|_1 \qquad \text{s.t. } Ax = y.$$

This problem has the Lagrangian

$$\mathcal{L}(x, y; \lambda) = \frac{1}{2}\|x - z\|_2^2 + \alpha\|y\|_1 - \lambda^T(Ax - y).$$

We obtain therefore the dual objective function

$$\begin{aligned} q(\lambda) &= \min_{x,y}\Big(\frac{1}{2}\|x - z\|_2^2 + \alpha\|y\|_1 - \lambda^T(Ax - y)\Big) \\ &= \min_{x}\Big(\frac{1}{2}\|x - z\|_2^2 - \lambda^T Ax\Big) + \min_{y}\big(\alpha\|y\|_1 + \lambda^T y\big). \end{aligned}$$

The minimum of the first term is attained for

$$x = z + A^T\lambda,$$

where we obtain the value

$$\begin{aligned} \min_{x}\Big(\frac{1}{2}\|x - z\|_2^2 - \lambda^T Ax\Big) &= \frac{1}{2}\|A^T\lambda\|_2^2 - \lambda^T A(z + A^T\lambda) \\ &= -\frac{1}{2}\|A^T\lambda\|_2^2 - z^T A^T\lambda \\ &= -\frac{1}{2}\|A^T\lambda + z\|_2^2 + \frac{1}{2}\|z\|_2^2. \end{aligned}$$

For the second term we have

$$\begin{aligned} \min_{y\in\mathbb{R}^m}\big(\alpha\|y\|_1 + \lambda^T y\big) &= \min_{y\in\mathbb{R}^m}\sum_{i=1}^{m}\big(\alpha|y_i| + \lambda_i y_i\big) \\ &= \sum_{i=1}^{m}\min_{t\in\mathbb{R}}\big(\alpha|t| + \lambda_i t\big), \end{aligned}$$

and

$$\min_{t\in\mathbb{R}}\big(\alpha|t| + \lambda_i t\big) = \begin{cases} -\infty & \text{if } |\lambda_i| > \alpha, \\ 0 & \text{if } |\lambda_i| \le \alpha. \end{cases}$$

Thus

$$\min_{y\in\mathbb{R}^m}\big(\alpha\|y\|_1 + \lambda^T y\big) = \begin{cases} -\infty & \text{if } \|\lambda\|_\infty > \alpha, \\ 0 & \text{if } \|\lambda\|_\infty \le \alpha. \end{cases}$$

As a consequence,

$$q(\lambda) = \begin{cases} -\frac{1}{2}\|A^T\lambda + z\|_2^2 + \frac{1}{2}\|z\|_2^2 & \text{if } \|\lambda\|_\infty \le \alpha, \\ -\infty & \text{if } \|\lambda\|_\infty > \alpha. \end{cases}$$

Thus the dual problem can be written as the constrained optimisation problem

$$(D_1) \qquad \max_{\lambda} -\frac{1}{2}\|A^T\lambda + z\|_2^2 + \frac{1}{2}\|z\|_2^2 \qquad \text{s.t. } \|\lambda\|_\infty \le \alpha.$$

The primal problem $(\hat{P}_1)$ is convex and the constraints are linear and feasible (the pair $(x, y) = (0, 0)$ is obviously feasible) and thus Theorem 12 is applicable. Since the primal objective function is coercive, it follows that there exists a primal solution, and consequently also a primal-dual solution (x^*, y^*, λ^*).

Moreover, if λ^* solves the dual problem (D_1), then (x^*, y^*) is a solution of

$$\min_{x,y}\mathcal{L}(x, y, \lambda^*)$$

(this follows from the fact that (x^*, y^*, λ^*) is a saddle point of the Lagrangian). As we have seen above during the computation of the dual objective function, the x-coordinate of the solution of this problem is $z + A^T\lambda^*$. Thus we conclude that

$$x^* = z + A^T\lambda^*,$$

where λ^* solves the dual problem

$$(D_1) \qquad \max_{\lambda} -\frac{1}{2}\|A^T\lambda + z\|_2^2 \qquad \text{s.t. } \|\lambda\|_\infty \le \alpha.$$

(Here we have omitted the constant term $\frac{1}{2}\|z\|_2^2$, as it only changes the value at the optimum, but not its position.)

In order to solve the dual optimisation problem, we can now use a projected gradient ascent method of the form

$$\hat{\lambda}^{(k+1)} = \lambda^{(k)} - \tau A(A^T\lambda^{(k)} + z),$$
$$\lambda^{(k+1)} = \pi_\Omega(\hat{\lambda}^{(k+1)}),$$

where π_Ω denotes the projection onto the (dual feasible) set

$$\Omega = \{\lambda \in \mathbb{R}^m : \|\lambda\|_\infty \le \alpha\}.$$

Explicitly, we have

$$\pi_\Omega(\lambda) = \min\{\alpha, \max\{-\alpha, \lambda\}\}.$$

As seen in the exercises, this projected gradient method converges provided the step length τ satisfies $2/\sigma_{\max}^2$, where $\sigma_{\max}$ denotes the largest singular value of A. The resulting method is summarised in Algorithm 1.

Initialisation: Choose a step length $0 < \tau < 2/\sigma_{\max}^2$ and $\lambda \in \mathbb{R}^m$;

Set $w \leftarrow \tau Az$;
while λ *not satisfactory* **do**
  $\lambda \leftarrow \lambda - \tau AA^T\lambda - w$;
  $\lambda \leftarrow \min\{\alpha, \max\{-\alpha, \lambda\}\}$;
end
$x^* = z + A^T\lambda$;

Algorithm 1: Dual projected gradient descent method for the solution of (P_1).

References

[1] Osman Güler. *Foundations of optimization*, volume 258 of *Graduate Texts in Mathematics*. Springer, New York, 2010.

[2] J. Nocedal and S. J. Wright. *Numerical optimization.* Springer Series in Operations Research and Financial Engineering. Springer, New York, second edition, 2006.

Department of Mathematics, Norwegian University of Science and Technology, 7491 Trondheim, Norway

E-mail address: markus.grasmair@ntnu.no