5.2 Nonlinear Programming

The aim of this section is to give an elementary discussion of nonlinear programming problems, i.e., finite dimensional optimization problems with a finite number of equality and inequality constraints.

The other chapters already contain a large number of results dealing with nonlinear programs, which were obtained as particular cases of results for more general optimization problems. We show in this section how some of these results can be obtained by using only standard differential calculus in $I\!R^n$ and the theory of finite dimensional linear programs. The section is mostly self contained; we use only a small number of results of other chapters whose proofs are based on elementary tools. Unless stated otherwise we assume that $x \cdot y := \sum_{i=1}^{n} x_i y_i$ denotes the standard scalar product of two vectors $x, y \in I\!R^n$, and that $\|x\| := (x \cdot x)^{1/2}$ denotes the Euclidean norm of a vector $x \in I\!R^n$. Specific results dealing with strong stability of strong regularity were obtained in section 5.1.6.

5.2.1 *Finite Dimensional Linear Programs*

In this section we discuss finite dimensional linear programming problems, i.e., optimization problems in finite dimensional spaces involving a linear objective function and finitely many linear constraints. In the theory of linear programming the following result is known as Farkas lemma.

Lemma 5.43 (Farkas lemma) *For any vectors $a_i \in I\!R^n$, $i = 0, \dots, p$, the following two conditions are equivalent:* (i) *for every $h \in I\!R^n$ such that $a_i \cdot h = 0$, $i = 1, \dots, q$, and $a_i \cdot h \le 0$, $i = q+1, \dots, p$, the inequality $a_0 \cdot h \ge 0$ holds,* (ii) *there exists $\lambda \in I\!R^p$, such that $a_0 + \sum_{i=1}^{p} \lambda_i a_i = 0$ and $\lambda_i \ge 0$ for all $i = q+1, \dots, p$.*

Proof. Rewriting an equality $a_i \cdot h = 0$ as two inequalities $a_i \cdot h \le 0$ and $-a_i \cdot h \le 0$, we can replace all equalitis in the above lemma by the corresponding linear inequalities. Therefore, we can assume without loss of generality that only the inequalities are present.

The implication (ii) $\Rightarrow$ (i) is immediate. Indeed, if h is such that $a_i \cdot h \le 0$, $i = 1, \dots, p$, then since $\lambda \in I\!R^p_+$, we have $a_0 \cdot h = -\sum_{i=1}^{p} \lambda_i a_i \cdot h \ge 0$.

We prove now that (i) implies (ii). By proposition 2.41, the convex set

$$E := \left\{ \sum_{i=1}^{p} \lambda_i a_i : \lambda_i \ge 0,\ i = 1, \dots, p \right\}$$

is closed. We need to prove that $-a_0 \in E$. Suppose that this is not true. Let $\bar{h}$ be the metric (orthogonal) projection of $-a_0$ onto E, i.e., $\bar{h}$ is the optimal solution of the optimization problem

$$\operatorname*{Min}_{h \in E} \|a_0 + h\|^2.$$

Note that this problem has a unique optimal solution, since E is nonempty, convex, and closed, and the objective function is continuous and strongly convex (lemma 2.33). Let $b := a_0 + \bar{h}$. Since $-a_0 \notin E$, we have that $b \neq 0$. Also, since E is convex, we have for any $e \in E$ and $t \in [0, 1]$ that $\bar{h} + t(e - \bar{h}) \in E$ and hence

$$b \cdot (e - \bar{h}) = \tfrac{1}{2} \lim_{t \downarrow 0} t^{-1} \left(\|a_0 + \bar{h} + t(e - \bar{h})\|^2 - \|a_0 + \bar{h}\|^2 \right) \geq 0.$$

Taking $e := 0$ we obtain that $b \cdot \bar{h} \leq 0$, and hence $b \cdot a_0 = b \cdot (b - \bar{h}) > 0$. On the other hand, taking $e := t^{-1} a_i$ for some $i \in \{1, \dots, p\}$, we get

$$0 \leq \lim_{t \downarrow 0} t \left(t^{-1} a_i - \bar{h} \right) \cdot b = b \cdot a_i.$$

By taking $h := -b$ in the condition (i), we obtain a contradiction. This completes the proof. □

Consider the following *linear programming* problem:

$$(LP) \qquad \begin{array}{ll} \operatorname{Min}_{x \in \mathbb{R}^n} & c \cdot x \\ \text{subject to} & a_i \cdot x = b_i, \ i = 1, \dots, q, \\ & a_i \cdot x \leq b_i, \ i = q+1, \dots, p. \end{array} \tag{5.91}$$

The dual of this problem can be written as follows:

$$(LD) \qquad \begin{array}{ll} \operatorname{Max}_{\lambda \in \mathbb{R}^p} & -b \cdot \lambda \\ \text{subject to} & c + \sum_{i=1}^p \lambda_i a_i = 0, \\ & \lambda_i \geq 0, \ i = q+1, \dots, p. \end{array} \tag{5.92}$$

Note that the primal problem (LP) can be written as the min-max problem

$$\operatorname*{Min}_{x \in \mathbb{R}^n} \left\{ \sup_{\lambda \in \mathbb{R}^p} L(x, \lambda) : \lambda_i \geq 0, \ i = q+1, \dots, p \right\}, \tag{5.93}$$

where

$$L(x, \lambda) := c \cdot x + \sum_{i=1}^p \lambda_i (a_i \cdot x - b_i)$$

is the corresponding Lagrangian function. By interchanging the minimization and maximization in (5.93), one obtains the corresponding dual problem (LD). Therefore, (LD) can be viewed as the Lagrangian dual of (LP).

Note also that the dual problem (LD) can be written in the format of the primal problem (LP). By writing then the dual of (LD) one obtains the primal problem (LP). Therefore, there is a complete symmetry here between the primal and its dual, and which one is called primal and which dual is somewhat arbitrary.

Theorem 5.44 (i) *Points $x \in \mathbb{R}^n$ and $\lambda \in \mathbb{R}^p$ are optimal solutions of (LP) and (LD), respectively, iff these points are feasible and the following complementarity* *condition holds:*

$$\lambda_i\,(a_i \cdot x - b_i) = 0, \quad i = q+1, \dots, p. \tag{5.94}$$

(ii) *Problems* (LP) *and* (LD) *have the same optimal value, unless* (LP) *and* (LD) *are both inconsistent. If their (common) optimal value is finite, then both problems have nonempty sets of optimal solutions.*

Proof. Let x and λ be feasible points of the problems (LP) and (LD), respectively. Since λ is a feasible point of (LD), the difference between the respective values of the objective functions can be written as

$$c \cdot x + b \cdot \lambda = -\left(\sum_{i=1}^{p} \lambda_i a_i\right) \cdot x + b \cdot \lambda = -\sum_{i=1}^{p} \lambda_i\,(a_i \cdot x - b_i). \tag{5.95}$$

Since x is feasible, and hence satisfies the constraints of (LP), and by the nonnegativity of the last components of λ, we obtain that the expression in the right hand side of (5.95) is nonnegative, and hence is equal to zero iff the complementarity condition holds. We see that for any feasible points x and λ of the respective problems, the corresponding value $c \cdot x$ of the primal problem is greater than or equal to the value $-b \cdot \lambda$ of the dual problem, and that equality holds iff the complementarity condition is satisfied. It follows that $\operatorname{val}(LP) \geq \operatorname{val}(LD)$, and if x and λ are feasible and the complementarity condition is satisfied, then x and λ are optimal.

Now, if both (LP) and (LD) are inconsistent, then the above assertions clearly hold. Therefore, we can assume that one of these problems, say, (LP), is consistent and hence $\operatorname{val}(LP) < +\infty$. Since $\operatorname{val}(LP) \geq \operatorname{val}(LD)$, we have that if $\operatorname{val}(LP) = -\infty$, then $\operatorname{val}(LD) = -\infty$, and hence in that case the assertions hold. Therefore, we can assume that (LP) has a finite optimal value. We have then by theorem 2.198 that (LP) has at least one optimal solution $\bar{x}$. Denote by

$$I(x) := \{i : a_i \cdot x = b_i,\ i = q+1, \dots, p\}$$

the set of inequality constraints active at x. If $h \in I\!R^n$ satisfies

$$a_i \cdot h = 0,\ i = 1, \dots, q; \quad a_i \cdot h \leq 0,\ i \in I(\bar{x}),$$

then for $t > 0$ small enough, we have that $\bar{x} + th$ is feasible for (LP). Since $\bar{x}$ is an optimal solution of (LP), it follows that $c \cdot \bar{x} \leq c \cdot (\bar{x} + th)$, and hence $c \cdot h \geq 0$. Then by the Farkas lemma (lemma 5.43), there exists $\lambda \in I\!R^p$ such that

$$c + \sum_{i=1}^{p} \lambda_i a_i = 0;\ \lambda_i \geq 0;\ q+1 \leq i \leq p;\ \lambda_i = 0,\ i \in \{q+1, \dots, p\} \setminus I(\bar{x}).$$

It follows that λ is a feasible point of (LD) and that the complementarity conditions hold. By (5.95) we obtain that λ is an optimal solution of (LD) and that $\operatorname{val}(LP) = \operatorname{val}(LD)$. Since there is no duality gap between the problems (LP) and (LD), it also follows from (5.95) that if x and λ are optimal, then the complementarity condition holds. This completes the proof. $\square$