Chapter 6

The Karush-Kuhn-Tucker conditions

6.1 Introduction

In this chapter we derive the first order necessary condition known as Karush-Kuhn-Tucker (KKT) conditions.

To this aim we introduce the *alternative theorems*.

6.2 Alternative theorems and the Farkas' Lemma

We present in this section *alternative theorems* which allows to formulate the *not existence* of solution as the existence of an alternative system.

Given two linear systems (I) e (II), an alternative theorem consists in proving that:

System (I) has a solution if and only if system (II) does not have a solution.

Among alternative theorems, one of the mst important that weare going to use to prove optimality conditions, is the *Farkas' Lemma*.

Theorem 6.3 (Farkas' Lemma) *Let B a $p \times n$ matrix and $g \in \Re^n$. System*

$$Bd \geq 0 \qquad g^T d < 0 \tag{I}$$

does not have a solution $d \in \Re^n$ *if and only if system*

$$B^T u = g \qquad u \geq 0 \tag{II}$$

has a solution. $u \in \Re^p$.

6.4 The Karush-Kuhn-Tucker conditions

Consider the problem

$$\begin{aligned} \min \quad & f(x) \\ & Ax \geq b \end{aligned} \qquad (P-POL)$$

In section 5 we derive the first order optimality conditions

If x^* is a local minimizer of problem (P-POL) then *there exists NO solution* $d \in R^n$ of the system

$$\begin{aligned} & A_{I(x^*)} d \geq 0, \\ & \nabla f(x^*)^T d < 0. \end{aligned} \qquad (6.1)$$

where $A_{I(x^*)}$ is the $|I(x^*)| \times n$ submatrix of A defined as $A_{I(x^*)} = (a_i^T)_{i \in I(x^*)}$ and $I(x^*) = \{i : \; a_i^T x^* = b_i\}$.

Identifying the system (6.1) with system (I) of Farkas' Lemma, we get

There exists NO solution $d \in R^n$ to the linear system (6.1) if and only if *there exists a solution* of the system

$$\begin{aligned} & A^T_{I(x^*)} u = \nabla f(x^*), \\ & u \geq 0. \end{aligned} \qquad (6.2)$$

where $u \in \mathbb{R}^{|I(x^*)|}$.

Using this result we can state the optimality conditions for problem (P-POL) known as *Karush-Kuhn-Tucker conditions* (KKT).

Theorem 6.5 (Karush-Kuhn-Tucker conditions) *If x^* is a local minimizer of problem (P-POL). Then a multiplier $\lambda^* \in R^m$ exists that such that*

(i) $\nabla f(x^*) - A^T \lambda^* = 0$,

(ii) $\lambda^* \geq 0$,

(iii) $\lambda_i^*(b_i - a_i^T x^*) = 0$ *for* $i = 1, \dots, m$.

(iv) $Ax^* \geq b$,

Proof. Since x^* is a local minimizer, by definition it is feasible hence (iv) holds. Further there exists no solution to the system (6.1). Identifying the system (6.1) with system (I) of Farkas' Lemma, we get that there exists NO solution che esiste to system(6.2).
Let $u_i^* \geq 0$, $i \in I(x^*)$ a solution of the preceding system and define $\lambda^* \in R^m$ as:

$$\lambda_i^* = \begin{cases} u_i^* & \text{for } i \in I(x^*) \\ 0 & \text{for } i \notin I(x^*). \end{cases} \qquad (6.3)$$

By definition $\lambda^* \geq 0$ and (ii) holds. Furthe (iii) derives immediately by the defintion (6.3) of λ^*. Indeed we have

$$\lambda_i^*(b_i - a_i^T x^*) = \begin{cases} u_i^*(b_i - a_i^T x^*) = 0 & \text{for } i \in I(x^*) \\ 0(b_i - a_i^T x^*) = 0 & \text{for } i \notin I(x^*). \end{cases}$$

Finally,(6.2) can be written as $\nabla f(x^*) = \sum_{i \in I(x^*)} u_i^* a_i$ so that

$$\nabla f(x^*) = \sum_{i \in I(x^*)} u_i^* a_i + \sum_{i \notin I(x^*)} 0 \cdot a_i = A^T \lambda^*$$

and (i) holds. □

> The vector $\lambda \in \Re^m$ is called *generalized Lagrange multiplier* or *KKT multiplier* associated to constraints $Ax \geq b$.

Condition (iii) is called *complementarity condition* and states that at optimality the product $\lambda_i^*(b_i - a_i^{*T} x^*)$ vanishes so that if the constraints is inactive the corresponding multiplier is zero. Since $\lambda^* \geq 0$ and $Ax^* - b \geq 0$ condition (iii)can be equivalently written as

$$\lambda^{*T}(b - Ax^*) = \sum_{i=1}^{m} \lambda_i^*(b_i - a_i^T x^*) = 0.$$

We can express conditions above using the *Lagrangian function.*

> The Lagrangian function associated to problem (P-POL) is
>
> $$L(x, \lambda) = f(x) + \lambda^T (b - Ax)$$

Condition (ii) states that in (x^*, λ^*) the gradient $\nabla_x L(x, \lambda) = \nabla f(x) - A^T \lambda$ vanishes, so that the pair (x^*, λ^*) ia a *stationary point* of the Lagrangian function.

> **(Karush-Kuhn-Tucker conditions)**
>
> Let x^* is a local minimizer of problem (P-POL). Then a multiplier $\lambda^* \in R^m$ exists that such that
>
> (i) $\nabla_x L(x^*, \lambda^*) = 0$ (stationary),
>
> (ii) $\lambda^* \geq 0$,
>
> (iii) $\lambda^{*T}(b - Ax^*) = 0$ (complementarity),
>
> (iv) $Ax^* \geq b$, (feasibility).

Candidates to be minimizer of problem (P-POL) can be found by solving the KKT conditions. We consider the following definition.

A point $\bar{x}$ is called *KKT point* or *stationary point* of problem (P-POL) if there exists $\bar{\lambda} \in \Re^m$ such that

(i) $A\bar{x} \geq b$,

(ii) $\nabla L(\bar{x}, \bar{\lambda}) = 0$,

(iii) $\bar{\lambda} \geq 0$,

(iv) $\bar{\lambda}^T(b - A\bar{x}) = 0$.

Under convexity assumption on f we get

Theorem 6.6 (Necessary and sufficient conditions for global optimzality) *Let f be a continuously differentiable convex function in $\Re^n$. A point x^* is a global minimizer of problem (P-POL) if and only if a multiplier $\lambda^* \in R^m$ exists that such that*

(i) $Ax^* \geq b$,

(ii) $\nabla f(x^*) - A^T \lambda^* = 0$,

(iii) $\lambda^* \geq 0$,

(iv) $\lambda^{*T}(b - Ax^*) = 0$.

If f is strictly convex and the conditions (i)-(iv) hold, then x^* is the unique global minimizer of f on S.

We derive now the KKT conditions for a more general problem.

$$\begin{array}{ll} \min & f(x) \\ & Dx = h \\ & Ax \geq b \end{array} \qquad (\mathrm{P-GEN})$$

where D is a $p \times n$ matrix, A is a $q \times n$ matrix, $h \in \Re^p$, $b \in \Re^q$.

The Lagrangian function associated to (P-GEN) is

$$L(x, \lambda, \mu) = f(x) + \lambda^T(b - Ax) + \mu^T(Dx - h),$$

with $\lambda \in \Re^q$, $\mu \in \Re^p$.

The KKT conditions for problem (P-GEN) are satisfied in $(x^*, \lambda^*, \mu^*) \in \Re^n \times \Re^q \times \Re^p$ when it holds:

(i) $Ax^* \geq b$, $Dx^* = h$ (feasibility),

(ii) $\nabla_x L(x^*, \lambda^*, \mu^*) = \nabla f(x^*) - A^T \lambda^* + D^T \mu^* = 0$ (stationarity),

(iii) $\lambda^* \geq 0$,

(iv) $\lambda^{*T}(b - Ax^*) = 0$ (complementarity).

Theorem 6.7 (Karush-Kuhn-Tucker Conditions for (P-GEN)) *Let x^* is a local minimizer of problem (P-GEN). Then multipliers $\lambda^* \in R^q$, $\mu^* \in \Re^p$ exist such that:*

(i) $\nabla_x L(x^*, \lambda^*, \mu^*) = 0$,

(ii) $\lambda^* \geq 0$,

(iii) $\lambda_i^*(b_i - a_i^T x^*) = 0$ *per* $i = 1, \dots, q$.

Consider the special case of *Quadratic Programming (QP) problems*

$$min \quad \frac{1}{2} x^T Q x + c^T x, \quad \text{(PQ)}$$

$$Ax \geq b$$

where Q is a $n \times n$ symetric and positive semidefinite matrix, $c \in \mathbb{R}^n$.

Theorem 6.8 (Optimality conditions for QP) *Let Q be a $n \times n$ symetric and positive semidefinite matrix. A point x^* is a global minimizer of problem* (PQ) *if and only if a multiplier $\lambda^* \in R^m$ exists that such that:*

(i) $Ax^* \geq b$,

(ii) $Qx^* + c - A^T \lambda^* = 0$,

(iii) $\lambda^* \geq 0$,

(iv) $\lambda^{*T}(b - Ax^*) = 0$.

Further if Q is positive definite and (i)-(iv) hold, then x^* is the unique global solution of problem (PQ).