Numerical Optimization
Lecture 17: KKT Conditions
Sangkyun Lee

The content is from Nocedal and Wright (2006) and Bertsekas (1999). Topics marked with ** are optional.

1 Constraint Qualification: LICQ

Constraint qualifications are the conditions to make sure \( T_\Omega(x) = \mathcal{F}(x) \). One of the simplest kind (but rather restrictive) is the linear independence constraint qualification (LICQ):

**Definition 17.1.** Given a point \( x \) and an active set \( A(x) \), we say LICQ holds if the set of active constraint gradients \( \{ \nabla c_i(x) : i \in A(x) \} \) is linearly independent.

We call \( x \) is regular if LICQ holds at \( x \).

**Lemma 17.2.** If the LICQ conditions hold at a feasible point \( x \), then \( \mathcal{F}(x) = T_\Omega(x) \).

**Proof.** ** We first define a matrix with active constraint gradients as rows, \( A(x^*)^T = [\nabla c_i(x^*)]_{i \in A(x^*)} \). Suppose that the matrix has \( m \) rows. Since LICQ holds, \( A(x^*) \) has full row rank \( m \). Let \( Z \) be a matrix whose columns are a basis for the null space of \( A(x^*) \), that is,

\[ Z \in \mathbb{R}^{n \times (n-m)}, \quad Z \text{ has full column rank}, \quad A(x^*)Z = 0. \]

Choose a vector \( d \in \mathcal{F}(x^*) \), and suppose that \( \{ t_k \} \) is any sequence of positive scalars satisfying \( \lim_{k \to \infty} t_k = 0 \). Define the parametrized system of equations \( R: \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}^n \) by

\[ R(z, t) = \begin{bmatrix} c(z) - tA(x^*)d \\ Z^T(z - x^* - td) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}. \]

We want to show that the solutions \( z = z_k \) of this system for small \( t = t_k > 0 \) give a feasible sequence approaching \( x^* \) and satisfying the definition of the tangent cone.

For \( t = 0 \), we set \( z = x^* \), and the Jacobian of \( R \) is then

\[ J_z R(x^*, 0) = \begin{bmatrix} A(x^*) \\ Z^T \end{bmatrix} \in \mathbb{R}^{n \times n}. \]

The Jacobian is nonsingular due to the construction of \( Z \). By the implicit function theorem, the system of equations above has a unique solution \( z_k \) for all sufficiently small values of \( t_k \). Moreover, from the system and the fact that \( d \in \mathcal{F}(x^*) \) we can check that

\[ i \in \mathcal{E} \Rightarrow c_i(z_k) = t_k \nabla c_i(x^*)^T d = 0, \]

\[ i \in A(x^*) \cap \mathcal{I} \Rightarrow c_i(z_k) = t_k \nabla c_i(x^*)^T d \geq 0, \]
and therefore $z_k$ is feasible. To check if this choice of $\{z_k\}$ satisfies the definition of the tangent cone, we use the fact that $R(z_k, t_k) = 0$ for all $k$ and Taylor’s theorem to find,

$$0 = R(z_k, t_k) = \begin{bmatrix} c(z_k) - t_k A(x^*) d \\ Z^T(z_k - x^* - t_k d) \end{bmatrix} = \begin{bmatrix} A(x^*) (z_k - x^*) + o(\|z_k - x^*\|_2) - t_k A(x^*) d \\ Z^T(z_k - x^* - t_k d) \end{bmatrix} = \begin{bmatrix} A(x^*) \\ Z^T \end{bmatrix} (z_k - x^* - t_k d) + o(\|z_k - x^*\|_2).$$

Dividing both sides by $t_k$ and using the invertibility of the matrix in the first term, we get

$$\frac{z_k - x^*}{t_k} = d + o(\|z_k - x^*\|_2/t_k).$$

That is, $d \in T_{\Omega^k}(x^*)$.

2 Lagrangian and the KKT Conditions

We define the Lagrangian function for general problems,

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

**Theorem 17.3.** Suppose that $x^*$ is a local minimizer, and $f$ and $c_i$ functions are continuously differentiable, and that LICQ holds at $x^*$ ($x^*$ is regular). Then there exists a unique Lagrange multiplier vector $\lambda^*$, with components $\lambda_i^*$, $i \in \mathcal{E} \cup \mathcal{I}$, such that the following conditions (a.k.a. the KKT conditions\(^1\)) are satisfied at $(x^*, \lambda^*)$,

\[
\begin{align*}
\nabla_x \mathcal{L}(x^*, \lambda^*) &= 0 \\
c_i(x^*) &= 0, \quad \text{for all } i \in \mathcal{E}, \\
c_i(x^*) &\geq 0, \quad \text{for all } i \in \mathcal{I}, \\
\lambda^*_i &\in \mathbb{R}, \quad \text{for all } i \in \mathcal{E}, \\
\lambda^*_i &\geq 0, \quad \text{for all } i \in \mathcal{I}, \\
\lambda^*_i c_i(x^*) &= 0, \quad \text{for all } i \in \mathcal{E} \cup \mathcal{I} \text{ (Complementarity Conditions)}
\end{align*}
\]

**Ex.**

$$\min_x \left( x_1 - \frac{3}{2} \right)^2 + \left( x_2 - \frac{1}{2} \right)^2, \quad \text{s.t.} \quad \begin{bmatrix} 1 & -x_1 & -x_2 \\ -x_1 & 1 + x_1 & x_2 \\ -x_2 & 1 + x_1 & 1 + x_1 \end{bmatrix} \geq 0.$$ 

The solution is at $x^* = (1, 0)^T$, and the first two constraints are active at this point. Denoting the two constraints by $c_1$ and $c_2$, we have

$$\nabla f(x^*) = \begin{bmatrix} -1 \\ -1/2 \end{bmatrix}, \quad \nabla c_1(x^*) = \begin{bmatrix} -1 \\ -1 \end{bmatrix}, \quad \nabla c_2(x^*) = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$$

We can check that the KKT conditions are satisfied with

$$\lambda^* = (3/4, 1/4, 0, 0)^T.$$

\(^1\)Karush-Kuhn-Tucker
References
