Necessary Conditions of General Nonlinear Constrained Optimization Problems and the Lagrange Multipliers

Kuei-Yuan Chan, chanky@mail.ncku.edu.tw
Professor of Mechanical Engineering,
National Taiwan University.

Let us consider continuous optimization problems of general nonlinear objective functions and nonlinear equality constraints in the form of Eq.(1).

$$\begin{aligned} \text{mimimize} f(\mathbf{x}) \\ \text{s.t.} \quad h_j(\mathbf{x}) = 0,\ j = 1 \cdots m \\ \forall \mathbf{x} \in \mathcal{X} \end{aligned} \tag{1}$$

We learned that the reduced gradient of Eq.(1) at the stationary point $\mathbf{x}_\dagger = (\mathbf{d}_\dagger, \mathbf{s}_\dagger)^T$ must satisfy

$$(\partial z / \partial \mathbf{d})_\dagger = \mathbf{0}^T$$

or one can also restate using the original symbols of Eq.(1) as

$$\left(\frac{\partial f}{\partial \mathbf{d}}\right)_\dagger - \left(\frac{\partial f}{\partial \mathbf{s}}\right)_\dagger \left(\frac{\partial \mathbf{h}}{\partial \mathbf{s}}\right)_\dagger^{-1} \left(\frac{\partial \mathbf{h}}{\partial \mathbf{d}}\right)_\dagger = \mathbf{0}^T \tag{2}$$

with all quantities being evaluated at $\mathbf{x}_\dagger$.

Let's define

$$\lambda^T \triangleq - \left(\frac{\partial f}{\partial \mathbf{s}}\right)_\dagger \left(\frac{\partial \mathbf{h}}{\partial \mathbf{s}}\right)_\dagger^{-1} \tag{3}$$

After rearrangements, we have substitute Eq.(3) into Eq.(2) as

$$\left(\frac{\partial f}{\partial \mathbf{d}}\right)_\dagger + \lambda^T \left(\frac{\partial \mathbf{h}}{\partial \mathbf{d}}\right)_\dagger = \mathbf{0}^T \tag{4}$$

Eq.(2) can also be

$$\left(\frac{\partial f}{\partial \mathbf{s}}\right)_\dagger + \lambda^T \left(\frac{\partial \mathbf{h}}{\partial \mathbf{s}}\right)_\dagger = \mathbf{0}^T \tag{5}$$

Combining Eqs.(4, 5), we have the *necessary condition* for a minimum : the gradient of the objective function must be the linear combination of the gradients of the constraints at the optimum as

$$\triangledown f(\mathbf{x}_\dagger) + \lambda^T \triangledown \mathbf{h}(\mathbf{x})_\dagger = \mathbf{0}^T \tag{6}$$

The stationary conditions in Eq.(6) is often expressed in terms of a special function, the Lagrangian function defined by

$$L(\mathbf{x}, \lambda) \triangleq f(\mathbf{x}) + \lambda^T \mathbf{h}(\mathbf{x}) \tag{7}$$

where λ) is called the **Lagrange multipliers**.

Let's extend the discussions to general problems with both equality and inequality constraints as

$$\begin{aligned} &\text{mimimize} f(\mathbf{x}) \\ \text{s.t.} \quad &\mathbf{h}(\mathbf{x}) = 0 \\ &\mathbf{g}(\mathbf{x}) \leq 0 \end{aligned} \tag{8}$$

At a constrained minimum, only some of the inequalities will be active. Let's represent the active inequalities as $\bar{\mathbf{g}}$. At the $\mathbf{x}_\star$, $\bar{\mathbf{g}}$ will be the same as $\mathbf{h}$. Therefore by extending Eq.(6), the first order optimality conditions will apply to the active inequalities

$$\bigtriangledown f(\mathbf{x}_\star) + \mu^T \bigtriangledown \bar{\mathbf{g}}(\mathbf{x}_\star) = \mathbf{0}^T \tag{9}$$

where μ^T is the Lagrange multipliers associated with the active inequalities.

The Karush-Kuhn-Tucker conditions is the first order necessary conditions for a problem with both equality and inequality constraints.

1. $\mathbf{h}(\mathbf{x}_\star) = 0,\ \mathbf{g}(\mathbf{x}_\star) \leq 0$;
2. $\bigtriangledown f_\star + \lambda^T \bigtriangledown \mathbf{h}_\star + \mu^T \bigtriangledown \mathbf{g}_\star = \mathbf{0}^T$, where $\lambda \neq \mathbf{0}, \mu \geq \mathbf{0}, \mu^T \mathbf{g} = \mathbf{0}$