Reduced Gradient for Equality Constrained Nonlinear Problems

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).

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

Any feasible point $\mathbf{x}$ must satisfy the constraints. Therefore a small perturbation $\partial\mathbf{x}$ about a feasible point will result in a perturbation ∂h_j. Apparently the point $\mathbf{x} + \partial\mathbf{x}$ is only feasible when $\partial h_j = 0, \forall j$.

The first order Taylor series approximation of perturbation for both objective functions and constraints are

$$\begin{aligned} \partial f &= \bigtriangledown f \partial\mathbf{x} = \sum_{i=1}^{n} (\partial f / \partial x_i)\, \partial x_i, \\ \partial h_j &= \bigtriangledown h_j \partial\mathbf{x} = \sum_{i=1}^{n} (\partial h_j / \partial x_i)\, \partial x_i = 0,\ j = 1, \cdots, m \end{aligned} \tag{2}$$

Eq.(2) has $(m+1)$ linear equations and $(n+1)$ unknowns. The final degrees of freedom is $(n-m)$. Let us define state variables s_i and decision variables d_i where

$$\mathbf{x} = [\mathbf{s}, \mathbf{d}] \tag{3}$$

where $s_i \triangleq x_i; i = 1, \cdots, m$ and $d_i \triangleq x_i; i = m+1, \cdots, n$.

When decision variables perturb with ∂d_i, the perturbation in the state variables must conform to feasibility as Eq.(4).

$$\begin{aligned} -\partial f + \sum_{i=1}^{m} (\partial f/\partial x_i)\partial x_i &= -\sum_{i=m+1}^{n} (\partial f/\partial x_i)\partial x_i \\ \sum_{i=1}^{m} (\partial h_j/\partial x_i)\partial x_i &= -\sum_{i=m+1}^{n} (\partial h_j/\partial x_i)\partial x_i,\ j = 1, \cdots, m \end{aligned} \tag{4}$$

Eq.(4) can be rewritten as

$$\begin{aligned} -\partial f + \sum_{i=1}^{m} (\partial f/\partial s_i)\partial s_i &= -\sum_{i=m+1}^{n} (\partial f/\partial d_i)\partial d_i \\ \sum_{i=1}^{m} (\partial h_j/\partial s_i)\partial s_i &= -\sum_{i=m+1}^{n} (\partial h_j/\partial d_i)\partial d_i,\ j = 1, \cdots, m \end{aligned} \tag{5}$$

Let's express Eq.(5) with the vector form as Eq.(6)

$$\begin{aligned} -\partial f + (\partial f/\partial \mathbf{s})\partial \mathbf{s} &= -(\partial f/\partial \mathbf{d})\partial \mathbf{d} \\ (\partial h_j/\partial \mathbf{s})\partial \mathbf{s} &= -(\partial h_j/\partial \mathbf{d})\partial \mathbf{d} \end{aligned} \tag{6}$$

Using the equality constraint perturbation in Eq.(6), we have

$$\partial \mathbf{s} = -(\partial \mathbf{h}/\partial (s))^{-1}(\partial \mathbf{h}/\partial \mathbf{d})\partial \mathbf{d}$$

and

$$\begin{aligned} \partial f &= (\partial f/\partial \mathbf{d})\partial \mathbf{d} + (\partial f/\partial \mathbf{s})\partial \mathbf{s} \\ & \left[(\partial f/\partial \mathbf{d}) - (\partial f/\partial \mathbf{s})(\partial \mathbf{h}/\partial \mathbf{s})^{-1}(\partial \mathbf{h}/\partial \mathbf{d})\right] \partial \mathbf{d} \end{aligned} \tag{7}$$

The quantity in the square bracket of Eq.(7) can be thought of as the gradient of a *new unconstrained* function $z(\mathbf{d})$, which will be equivalent to the original objective function f if the solution variables had been eliminated. Thus we can define a quantity

$$\partial z/\partial \mathbf{d} = (\partial f/\partial \mathbf{d}) - (\partial f/\partial \mathbf{s})(\partial \mathbf{h}/\partial \mathbf{s})^{-1}(\partial \mathbf{h}/\partial \mathbf{d}) \tag{8}$$

which we call **the reduced gradient** of the function f. The feasible domain of z is in the $(n-m)$ dimensional space; the function z is considered unconstrained since we assume $\mathbf{d}$ to be interior point. Thus, the obvious condition for a (constrained) stationary point $\mathbf{x}_\dagger = (\mathbf{d}_\dagger, \mathbf{s}_\dagger)^T$ is that

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