Nonlinear Programming Problem (NLP) (Version 3.0)

Karush-Kuhn-Tucker (KKT) Conditions

Consider a scalar cost L which is a function of the parameter vector $\mathbf{y}$ *in* R^p and the constraint vector $\mathbf{f}$ *in* R^q. The constraints are of the form $\mathbf{f}(\mathbf{y}) \leq \mathbf{0}$, by which is meant that each element of $\mathbf{f}$ is nonpositive.

A Lagrange multipler vector $\boldsymbol{\lambda}$ *in* R^q is introduced and the scalar H defined as

$$H(\mathbf{y}, \boldsymbol{\lambda}) \equiv L(\mathbf{y}) + \boldsymbol{\lambda}^T \mathbf{f}(\mathbf{y})$$

Thus there are a total of $p + q$ unknowns, namely the elements of $\mathbf{y}$ and $\boldsymbol{\lambda}$.

Necessary Conditions

$$\frac{\partial H}{\partial \mathbf{y}} = \mathbf{0}^T \qquad \text{(NC-1)}$$

$$\lambda_k \geq 0 \quad \text{if} \quad f_k(\mathbf{y}) = 0 \quad \text{(active)} \qquad \text{(NC-2a)}$$

$$\lambda_k = 0 \quad \text{if} \quad f_k(\mathbf{y}) < 0 \quad \text{(inactive)} \qquad \text{(NC-2b)}$$

Note that regardless of whether a constraint is active or inactive, the product $\lambda_k f_k = 0$, resulting in the fact that H is equal to L. Next, order the elements of the constraint vector so that the first n elements, where $0 \leq n \leq q$, correspond to the active constraints.
Thus in (NC-2a) $k = 1 \rightarrow n$ (if $n > 0$) and in (NC-2b) $k = n + 1 \rightarrow q$ (if $n < q$).

Note that (NC-1) provides p (scalar) equations and that (NC-2a) and (NC-2b) together provide q equations, because either $f_k = 0$ or $\lambda_k = 0$.
These equations can then be solved for the $p + q$ unknowns, the elements of $\mathbf{y}$ and $\boldsymbol{\lambda}$, to determine the constrained stationary point(s) of the function L.

It is convenient to partition the parameter vector $\mathbf{y}$ and the constraint vector $\mathbf{f}$ based on the number of active constraints, n. Let

$$\mathbf{y} = \begin{bmatrix} \mathbf{x} \\ \mathbf{u} \end{bmatrix}$$

with $\mathbf{x}$ *in* R^n and $\mathbf{u}$ *in* R^m, where, by definition $m \equiv p - n$, i.e., m is equal to the number of parameters, p, minus the number of active constraints, n. Thus m is the number of free parameters (degrees of freedom) that can be varied independently to minimize L.

Next, partition the constraint vector as:

$$\mathbf{f} = \begin{bmatrix} \boldsymbol{\alpha} \\ \boldsymbol{\beta} \end{bmatrix}$$

with $\boldsymbol{\alpha}$ *in* R^n and $\boldsymbol{\beta}$ *in* R^s, where $s = q - n$ is the number of inactive constraints. The selection of which elements of the parameter vector $\mathbf{y}$ comprise the vector $\mathbf{x}$ is arbitrary, except that the *constraint qualification* must be satisfied. This condition requires that the gradients of the active constraint functions with respect to the parameter vector must be linearly independent. This is necessary in order that (NC-1) be valid, i.e., that the gradient of the cost can be expressed as a linear combination of the active constraint gradients, with the (negatives of the) Lagrange multipliers as coefficients.

(over)

The constraint qualification will be satisfied if the $n \times n$ Jacobian matrix of the vector of active constraint functions $\boldsymbol{\alpha}$ with respect to the vector $\mathbf{x}$ is nonsingular:

$$\det \frac{\partial \boldsymbol{\alpha}}{\partial \mathbf{x}} \neq 0 \qquad \text{(NC-3)}$$

and the $m \times m$ Hessian matrix (see below)

$$\left(\frac{\partial^2 L}{\partial \mathbf{u}^2} \right)_{\boldsymbol{\alpha} = \mathbf{0}} \text{ is positive semidefinite.} \qquad \text{(NC-4)}$$

Sufficient Conditions

The sufficient conditions can then be stated as:

$$\lambda_k > 0 \quad , \quad k = 1 \rightarrow n \qquad \text{(SC-1)}$$

and the $m \times m$ Hessian matrix

$$\left(\frac{\partial^2 L}{\partial \mathbf{u}^2} \right)_{\boldsymbol{\alpha} = \mathbf{0}} \text{ is positive definite.} \qquad \text{(SC-2)}$$

An expression for this Hessian matrix is given by:

$$\left(\frac{\partial^2 L}{\partial \mathbf{u}^2} \right)_{\boldsymbol{\alpha} = \mathbf{0}} = H_{uu} + H_{ux} Z + (H_{ux} Z)^T + Z^T H_{xx} Z$$

where $Z \equiv - \boldsymbol{\alpha}_{\mathbf{x}}^{-1} \boldsymbol{\alpha}_{\mathbf{u}}$ and H_{ux} is the $m \times n$ matrix given by

$$H_{ux} = \frac{\partial}{\partial \mathbf{x}} \left(\frac{\partial H}{\partial \mathbf{u}} \right)^T$$

(SC-1) is a strengthened version of (NC-2a) and states that the cost can be decreased only by violating the $\boldsymbol{\alpha} \leq \mathbf{0}$ constraint, while (SC-2) states that the cost function being locally convex at the constrained stationary point indicates a local minimum. Note that this convexity condition pertains only to the free parameters comprising the vector $\mathbf{u}$ and that the Hessian matrix evaluated in (SC-2) is evaluated holding the active constraints $\boldsymbol{\alpha} = \mathbf{0}$.

The conditions summarized above form the Karush-Kuhn-Tucker (KKT) Conditions for the Nonlinear Programming Problem.

J. E. Prussing
sept 2004

(over)