INTERPRETATION OF LAGRANGE MULTIPLIERS

IN PARAMETER OPTIMIZATION

by

John E. Prussing

Consider the problem of minimizing $L(\mathbf{x}, \mathbf{u})$ subject to the n-dimensional constraint equation:

$$\mathbf{f}(\mathbf{x}, \mathbf{u}, \mathbf{c}) = \mathbf{0} \tag{1}$$

where $\mathbf{x}$ is n-dimensional, $\mathbf{u}$ is m-dimensional, and $\mathbf{c}$ is a q-dimensional vector of constants in the constraint equation.

The differential change in the vector $\mathbf{f}$ is:

$$d\mathbf{f} = \mathbf{f}_x \, d\mathbf{x} + \mathbf{f}_u \, d\mathbf{u} + \mathbf{f}_c \, d\mathbf{c} \tag{2}$$

The matrix $\mathbf{f}_c$ is an $n \times q$ Jacobian matrix. The first two terms in eqn (2) are what are called "$d\mathbf{f}$" in Ref. 1, eqns. (1.2.4) and (1.4.3).

Considering differential changes in variables from their optimal values, one wishes to hold the constraint $\mathbf{f} = \mathbf{0}$ by requiring $d\mathbf{f} = \mathbf{0}$. This results in

$$d\mathbf{x} = -\mathbf{f}_x^{-1} (\mathbf{f}_u \, d\mathbf{u} + \mathbf{f}_c \, d\mathbf{c}) \tag{3}$$

Now,

$$dL = L_x \, d\mathbf{x} + L_u \, d\mathbf{u} \tag{4}$$

which is equal to

$$dL = (L_u - L_x \, \mathbf{f}_x^{-1} \, \mathbf{f}_{u)} \, d\mathbf{u} - L_x \, \mathbf{f}_x^{-1} \, \mathbf{f}_c \, d\mathbf{c} \tag{5}$$

From the NC that $H_x = \mathbf{0}^T$, $\boldsymbol{\lambda}^T = -L_x \, \mathbf{f}_x^{-1}$, which, along with $H_u = \mathbf{0}^T$ yields:

$$L_u - L_x \, \mathbf{f}_x^{-1} \, \mathbf{f}_u = \mathbf{0}^T = \left(\frac{\partial L}{\partial \mathbf{u}}\right)_{\mathbf{f} = \mathbf{0}} \tag{6}$$

Substituting into eqn (5) yields:

$$dL = \boldsymbol{\lambda}^T \, \mathbf{f}_c \, d\mathbf{c} \tag{7}$$

and thus,

$$\frac{\partial L}{\partial \mathbf{c}} = \boldsymbol{\lambda}^T \, \mathbf{f}_c \tag{8}$$

should replace eqn. (1.4.9) of Ref. 1.

The gradient $\frac{\partial L}{\partial \mathbf{c}}$ is an $1 \times q$ row vector. Each element represents the *sensitivity* of the minimum value of the cost L to a change in the corresponding component of the vector $\mathbf{c}$, i.e., the change in a particular constant in the constraint equation.

In component form,

$$dL = \lambda_j \frac{\partial f_j}{\partial c_k} dc_k \tag{9}$$

or,

$$\frac{\partial L}{\partial c_k} = \lambda_j \frac{\partial f_j}{\partial c_k} \tag{10}$$

Note that all values above are evaluated at the optimum.

Reference

1. Bryson, A.E. Jr. and Ho, Y-C, *Applied Optimal Control*, Hemisphere Publishing Corp., 1975,