Lagrange multipliers and sensitivity

Mark S. Gockenbach

The Lagrange multipliers for the NLP

\[
\min f(x) \quad \text{s.t.} \quad g(x) = 0
\]

have an interpretation that I have not mentioned yet: In a certain sense, they quantify the change in the optimal value of (1) if the constraints are perturbed. To be precise, I define

\[ p(u) = \min \{ f(x) : g(x) = u \}. \]

Provided \( x^* \) is a nonsingular point, the function \( p \) is defined on some neighborhood \( N \) of 0 in \( \mathbb{R}^m \). Moreover, there exist smooth functions \( x : N \to \mathbb{R}^n \) and \( \lambda : N \to \mathbb{R}^m \) such that \( x(u) \) is the (locally) unique solution to the NLP defining \( p \) and \( \lambda(u) \) is the corresponding Lagrange multiplier.\(^1\)

Therefore

\[ p(u) = f(x(u)) \]

and

\[ g(x(u)) = u. \]

The desired result follows from differentiating (2) and (3) and combining the results. By the chain rule, differentiating (2) yields

\[ \nabla p(u) = \nabla x(u) \nabla f(x(u)) \]

and differentiating (3) yields

\[ \nabla x(u) \nabla g(x(u)) = I. \]

In particular,

\[
\begin{align*}
\nabla p(0) &= \nabla x(0) \nabla f(x^*), \\
\nabla x(0) \nabla g(x^*) &= I.
\end{align*}
\]

Since \( \nabla f(x^*) = \nabla g(x^*) \lambda^* \), the first equation yields

\[
\begin{align*}
\nabla p(0) &= \nabla x(0) (\nabla g(x^*) \lambda^*) \\
&= (\nabla x(0) \nabla g(x^*)) \lambda^* \\
&= I \lambda^* \\
&= \lambda^*.
\end{align*}
\]

This is the desired result:

\[ \nabla p(0) = \lambda^*. \]

\(^1\)This can be proved by applying the implicit function theorem to the system

\[
\begin{align*}
\nabla f(x) - \nabla g(x) \lambda &= 0, \\
-g(x) + u &= 0.
\end{align*}
\]
This shows that if the original constraints \( g(x) = 0 \) are perturbed to \( g(x) = u \), where \( u \) is small, then the optimal value of the NLP is changed to approximately

\[
\nabla p(0) \cdot u = \lambda^* \cdot u = \sum_{i=1}^{m} \lambda_i^* u_i.
\]

**Example 0.1** As an explicit example of the above results, I will solve

\[
\min_{s.t.} f(x) \quad \text{subject to } g(x) = u,
\]

where \( f : \mathbb{R}^3 \to \mathbb{R} \) and \( g : \mathbb{R}^3 \to \mathbb{R} \) are defined by \( f(x) = x_1 + x_2 + x_3 \) and \( g(x) = x_1^2 + x_2^2 + x_3^2 - 1 \). It is easy to show that the global minimizer of the NLP is

\[
x(\mu) = -\frac{\sqrt{1 + u}}{\sqrt{3}} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}
\]

and the corresponding Lagrange multiplier is

\[
\lambda(\mu) = -\frac{\sqrt{3}}{2\sqrt{1 + u}}
\]

(provided \( \mu > -1 \)). In particular, with \( u = 0 \), the solution and Lagrange multiplier are

\[
x^* = -\frac{1}{\sqrt{3}} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \quad \lambda^* = -\frac{\sqrt{3}}{2}
\]

Then

\[
p(u) = f(x(\mu)) = -\sqrt{3}\sqrt{1 + u},
\]

so

\[
p'(u) = -\frac{\sqrt{3}}{2\sqrt{1 + u}}
\]

and, in particular,

\[
p'(0) = -\frac{\sqrt{3}}{2} = \lambda^*,
\]

as predicted.