A new proof of the Lagrange multiplier rule
Jan Brinkhuis and Vladimir Protassov

Abstract. We present an elementary self-contained proof for the Lagrange multiplier rule. It does not refer to any substantial preparations and it is only based on the observation that a certain limit is positive. At the end of this note, the power of the Lagrange multiplier rule is analyzed.

Keywords. Lagrange multiplier rule; penalty function.

We prove the following.

Theorem (the Lagrange multiplier rule). Assume that $G$ is a set in $\mathbb{R}^n$ and that $f_i, 0 \leq i \leq m$ are continuously differentiable functions on $G$. Let $\hat{x}$ be an interior point of $G$ that is a minimizer of $f_0(x)$ on the set of of points in $G$ for which $f_i(x) = 0, 1 \leq i \leq m$. Then the set of derivatives $f'_i(\hat{x}), 0 \leq i \leq m$, the $n$-dimensional row vectors of the partial derivatives of the $f_i, 0 \leq i \leq m$ in $\hat{x}$, is linearly dependent.

The given formulation of the conclusion of the rule is equivalent to the following one, which is more familiar: if the derivatives $f'_i(\hat{x}), 1 \leq i \leq m$ are linearly independent, then $\hat{x}$ is a stationary point of the function $f_0(x)+\lambda_1 f_1(x)+\cdots+\lambda_m f_m(x)$ for suitable numbers (Lagrange multipliers) $\lambda_i, 1 \leq i \leq m$.

The existing proofs of the multiplier rule in the literature require effort. Either they refer to substantial preparations or they contain technical arguments. Such preparations are the implicit function theorem ([1],[2]), the inverse function theorem (of which the multiplier rule is an immediate consequence [3]), the tangent space theorem ([4], [5]), the Brouwer fixed-point theorem ([6]) or the Ekeland variational principle ([7]). Technical arguments are needed to establish inequalities and estimates ([8]). Then the resulting proofs are elementary but they tend to be relatively long and involved ([9], 10]). The new proof given in the present note requires essentially no effort: it is not only elementary, but also simple, short and insightful. Its novelty consists in the easy observations that a certain limit is positive and that this implies that boundary points are excluded as minimizers of a certain penalty function. The remainder of the proof is obvious. All elementary proofs for the multiplier rule, including the one given in this note, make use of the theorem that a continuous function assumes its minimal value over a closed ball.

Proof. By contraposition.

• Preparations. Assume that $\hat{x}$ is an interior point of $G$ that is a solution of the system of equations $f_i(x) = 0, 1 \leq i \leq m$ for which $f'_i(\hat{x}), 0 \leq i \leq m$ are linearly independent. We will show that $\hat{x}$ is not a minimum of the function $f_0(x)$ on the set of $x \in G$ for which $f_i(x) = 0, 1 \leq i \leq m$.

Without loss of generality it is assumed that $f_0(\hat{x}) = 0$, $\hat{x} = 0$ and $n = m + 1$—the latter can be achieved by adding if necessary constraints $a_i x = 0, i \in I$, chosen in such a way that the set...
Here is a precise analytic verification. The stationarity vector equation
\[ F(x) = F'(0)x + o(|x|)(x \to 0) \]
(1)

(‘Landau little o’) as the \( f_i, 0 \leq i \leq m \) are continuously differentiable and \( F(0) = 0 \). The assumption that \( f_i'(0), 0 \leq i \leq m = n - 1 \) are linearly independent means that the \( n \times n \)-matrix \( F'(0) \) has rank \( m + 1 = n \), so \( F'(0) \) is invertible. Choose for each \( k \in \mathbb{N} \) a minimizer \( x_k \) of the penalty function
\[ p_k(x) = |F(x) + (1/k^2, 0, \ldots, 0)^T| \]
on the closed ball \(|x| \leq 1/k\), where \(| \cdot |\) is the Euclidean norm.

- **Exclusion of boundary minimizers for the penalty function.** We consider the limit
\[ L = \lim_{x \to 0} \frac{|F(x) + (|x|^2, 0, \ldots, 0)^T - (|x|^2, 0, \ldots, 0)^T|}{|F'(0)x|} \]
(3)
Here we recall that \( F'(0) \) is invertible. Note that
\[ \lim_{x \to 0} \frac{|x|^2}{|F'(0)x|} = 0. \]
Therefore,
\[ L = \lim_{x \to 0} \frac{|F(x)|}{|F'(0)x|} = 1 > 0, \]
(4)
using (1) and the invertibility of \( F'(0) \). It follows from (3) and (4) that for \( k \) sufficiently large and for each point \( \bar{x} \) on the sphere \(|x| = 1/k\), one has
\[ |F'(0)\bar{x} + (|\bar{x}|^2, 0, \ldots, 0)^T| > |(|\bar{x}|^2, 0, \ldots, 0)^T|, \]
that is, the penalty function \( p_k \), given by (2), takes a larger value in \( \bar{x} \) than in \( 0 \)—here we use also \( F(0) = 0 \). Therefore, boundary points of the ball \(|x| \leq 1/k\) cannot be minimizers of the penalty function \( p_k \), given by (2), on this ball. That is, \( x_k \) is an interior point of this ball, \(|x_k| < 1/k\).

- **Conclusion of the proof.** Moreover, for \( k \) sufficiently large, \(|x_k| < 1/k\) implies that the matrix \( F'(x_k) \) is invertible as \( F'(0) \) is invertible and \( F' \) is continuous. From this, and as the minimizer \( x_k \) is an interior point of the closed ball \(|x| \leq 1/k\), it is obvious that
\[ F(x_k) + (1/k^2, 0, \ldots, 0)^T = 0. \]
(5)
Indeed, otherwise a local decrease of the penalty function \( p_k \), given by (2), would be possible. Here is a precise analytic verification. The stationarity vector equation
\[ 0 = \frac{d}{dx} |F(x) + (1/k^2, 0, \ldots, 0)^T|^2 |x = x_k = 2(F(x_k)^T + (1/k^2, 0, \ldots, 0))F'(x_k) \]
holds as the minimizer $x_k$ is interior. Here we use that the minimizers of a nonnegative valued function do not change if the objective function is squared and that $|v|^2 = v^T v$ for a column vector $v$. This gives (5) as $F'(x_k)$ is invertible. That is, by the definition of $F$,

$$f_0(x_k) = -1/k^2 < 0 = f_0(0), f_i(x_k) = 0, 1 \leq i \leq m.$$ 

It follows that $\widehat{x} = 0$ is not a minimizer of the function $f_0(x)$ on the set of points in $G$ for which $f_i(x) = 0, 1 \leq i \leq m$. This concludes the proof of the theorem.

**Comment on the power of the rule.** The power of the rule lies in the reversal of the natural order of two main tasks, elimination and differentiation. The natural order would be to eliminate $m$ of the variables $x_1, \ldots, x_n$, using the equality constraints $f_i(x) = 0, 1 \leq i \leq m$, substitute into the objective function $f_0(x)$, and then put the partial derivatives of the resulting expression in $n - m$ variables equal to zero. The elimination task is often a nonlinear problem that is difficult or even impossible. Differentiating first gives the following statement, which is a reformulation of the conclusion of the theorem: all solutions $h \in \mathbb{R}^n$ of the system of linear equations

$$f_i'(\widehat{x})h = 0, 1 \leq i \leq m \tag{6}$$

satisfy

$$f_0'(\widehat{x})h = 0,$$

provided the vectors $f_i'(\widehat{x}), 1 \leq i \leq m$ are linearly independent. The remaining elimination task is now just a linear problem. It is easy to eliminate $m$ of the $n$ variables $h_1, \ldots, h_n$ using the system of linear equations (6). Then one can substitute in $f_0'(\widehat{x})h$. Finally one should put the coefficients of the resulting expression in $n - m$ variables equal to zero. This completes elimination without multipliers. In particular, multipliers are not essential to the power of the rule, but merely provide a convenient way to carry out the linear elimination task. See Ch 3 of [4] for more details on the power of the rule.

**Acknowledgment.** We would like to thank Jan van de Craats for his suggestions on the presentation.

**References.**


