Because by assumption $x^*$ is a regular point and $L(x^*)$ is positive definite on $M$, it follows that this matrix is nonsingular (see Exercise 11). Thus, by the Implicit Function Theorem, there is a solution $x(c), \lambda(c)$ to the system which is in fact continuously differentiable.

By the chain rule we have

$$\nabla_c f(x(c)) \bigg|_{c=0} = \nabla_x f(x^*) \nabla_c x(0).$$

and

$$\nabla_c h(x(c)) \bigg|_{c=0} = \nabla_x h(x^*) \nabla_c x(0).$$

In view of (31), the second of these is equal to the identity $I$ on $E^m$, while this, in view of (30), implies that the first can be written

$$\nabla_c f(x(c)) \bigg|_{c=0} = -\lambda^T. \blacksquare$$

## 11.8 INEQUALITY CONSTRAINTS

We consider now problems of the form

$$\begin{align*}
\text{minimize} & \quad f(x) \\
\text{subject to} & \quad h(x) = 0 \\
& \quad g(x) \leq 0.
\end{align*} \tag{32}$$

We assume that $f$ and $h$ are as before and that $g$ is a $p$-dimensional function. Initially, we assume $f, h, g \in C^1$.

There are a number of distinct theories concerning this problem, based on various regularity conditions or constraint qualifications, which are directed toward obtaining definitive general statements of necessary and sufficient conditions. One can by no means pretend that all such results can be obtained as minor extensions of the theory for problems having equality constraints only. To date, however, these alternative results concerning necessary conditions have been of isolated theoretical interest only—for they have not had an influence on the development of algorithms, and have not contributed to the theory of algorithms. Their use has been limited to small-scale programming problems of two or three variables. We therefore choose to emphasize the simplicity of incorporating inequalities rather than the possible complexities, not only for ease of presentation and insight, but also because it is this viewpoint that forms the basis for work beyond that of obtaining necessary conditions.
First-Order Necessary Conditions

With the following generalization of our previous definition it is possible to parallel the development of necessary conditions for equality constraints.

**Definition.** Let \( x^* \) be a point satisfying the constraints

\[
    h(x^*) = 0, \quad g(x^*) \leq 0, \tag{33}
\]

and let \( J \) be the set of indices \( j \) for which \( g_j(x^*) = 0 \). Then \( x^* \) is said to be a *regular point* of the constraints (33) if the gradient vectors \( \nabla h_i(x^*), \nabla g_j(x^*), 1 \leq i \leq m, j \in J \) are linearly independent.

We note that, following the definition of active constraints given in Section 11.1, a point \( x^* \) is a regular point if the gradients of the active constraints are linearly independent. Or, equivalently, \( x^* \) is regular for the constraints if it is regular in the sense of the earlier definition for equality constraints applied to the active constraints.

**Karush–Kuhn–Tucker Conditions.** Let \( x^* \) be a relative minimum point for the problem

\[
    \text{minimize } f(x) \\
    \text{subject to } h(x) = 0, \quad g(x) \leq 0, \tag{34}
\]

and suppose \( x^* \) is a regular point for the constraints. Then there is a vector \( \lambda \in E^m \) and a vector \( \mu \in E^p \) with \( \mu \geq 0 \) such that

\[
    \nabla f(x^*) + \lambda^T \nabla h(x^*) + \mu^T \nabla g(x^*) = 0 \tag{35}
\]

\[
    \mu^T g(x^*) = 0. \tag{36}
\]

**Proof.** We note first, since \( \mu \geq 0 \) and \( g(x^*) \leq 0 \), (36) is equivalent to the statement that a component of \( \mu \) may be nonzero only if the corresponding constraint is active. This a *complementary slackness* condition, stating that \( g(x^*), < 0 \) implies \( \mu_i = 0 \) and \( \mu_i > 0 \) implies \( g(x^*)_i = 0 \).

Since \( x^* \) is a relative minimum point over the constraint set, it is also a relative minimum over the subset of that set defined by setting the active constraints to zero. Thus, for the resulting equality constrained problem defined in a neighborhood of \( x^* \), there are Lagrange multipliers. Therefore, we conclude that (35) holds with \( \mu_j = 0 \) if \( g_j(x^*) \neq 0 \) (and hence (36) also holds).

It remains to be shown that \( \mu \geq 0 \). Suppose \( \mu_k < 0 \) for some \( k \in J \). Let \( S \) and \( M \) be the surface and tangent plane, respectively, defined by all other active constraints at \( x^* \). By the regularity assumption, there is a \( y \) such that \( y \in M \) and
$\nabla g_k(x^*)y < 0$. Let $x(t)$ be a curve on $S$ passing through $x^*$ (at $t = 0$) with $\dot{x}(0) = y$. Then for small $t \geq 0$, $x(t)$ is feasible, and

$$\left. \frac{df}{dt}(x(t)) \right|_{t=0} = \nabla f(x^*)y < 0$$

by (35), which contradicts the minimality of $x^*$. □

**Example.** Consider the problem

minimize $2x_1^2 + 2x_1 x_2 + x_2^2 - 10x_1 - 10x_2$

subject to $x_1^2 + x_2^2 \leq 5$

$3x_1 + x_2 \leq 6$.

The first-order necessary conditions, in addition to the constraints, are

$$4x_1 + 2x_2 - 10 + 2\mu_1 x_1 + 3\mu_2 = 0$$

$$2x_1 + 2x_2 - 10 + 2\mu_1 x_2 + \mu_2 = 0$$

$$\mu_1 \geq 0, \quad \mu_2 \geq 0$$

$$\mu_1 (x_1^2 + x_2^2 - 5) = 0$$

$$\mu_2 (3x_1 + x_2 - 6) = 0.$$ 

To find a solution we define various combinations of active constraints and check the signs of the resulting Lagrange multipliers. In this problem we can try setting none, one, or two constraints active. Assuming the first constraint is active and the second is inactive yields the equations

$$4x_1 + 2x_2 - 10 + 2\mu_1 x_1 = 0$$

$$2x_1 + 2x_2 - 10 + 2\mu_1 x_2 = 0$$

$$x_1^2 + x_2^2 = 5,$$

which has the solution

$$x_1 = 1, \quad x_2 = 2, \quad \mu_1 = 1.$$ 

This yields $3x_1 + x_2 = 5$ and hence the second constraint is satisfied. Thus, since $\mu_1 > 0$, we conclude that this solution satisfies the first-order necessary conditions.
Second-Order Conditions

The second-order conditions, both necessary and sufficient, for problems with inequality constraints, are derived essentially by consideration only of the equality constrained problem that is implied by the active constraints. The appropriate tangent plane for these problems is the plane tangent to the active constraints.

**Second-Order Necessary Conditions.** Suppose the functions $f, g, h \in C^2$ and that $x^*$ is a regular point of the constraints (33). If $x^*$ is a relative minimum point for problem (32), then there is a $\lambda \in E^m$, $\mu \in E^p$, $\mu \geq 0$ such that (35) and (36) hold and such that

$$L(x^*) = F(x^*) + \lambda^T H(x^*) + \mu^T G(x^*)$$

is positive semidefinite on the tangent subspace of the active constraints at $x^*$.

**Proof.** If $x^*$ is a relative minimum point over the constraints (33), it is also a relative minimum point for the problem with the active constraints taken as equality constraints.

Just as in the theory of unconstrained minimization, it is possible to formulate a converse to the Second-Order Necessary Condition Theorem and thereby obtain a Second-Order Sufficiency Condition Theorem. By analogy with the unconstrained situation, one can guess that the required hypothesis is that $L(x^*)$ be positive definite on the tangent plane $M$. This is indeed sufficient in most situations. However, if there are degenerate inequality constraints (that is, active inequality constraints having zero as associated Lagrange multiplier), we must require $L(x^*)$ to be positive definite on a subspace that is larger than $M$.

**Second-Order Sufficiency Conditions.** Let $f, g, h \in C^2$. Sufficient conditions that a point $x^*$ satisfying (33) be a strict relative minimum point of problem (32) is that there exist $\lambda \in E^m$, $\mu \in E^p$, such that

$$\mu \geq 0$$

$$\mu^T g(x^*) = 0$$

$$\nabla f(x^*) + \lambda^T \nabla h(x^*) + \mu^T \nabla g(x^*) = 0,$$

and the Hessian matrix

$$L(x^*) = F(x^*) + \lambda^T H(x^*) + \mu^T G(x^*)$$

is positive definite on the subspace

$$M' = \{ y : \nabla h(x^*)y = 0, \nabla g_j(x^*)y = 0 \text{ for all } j \in J \},$$

where

$$J = \{ j : g_j(x^*) = 0, \mu_j > 0 \}. $$
Proof. As in the proof of the corresponding theorem for equality constraints in Section 11.5, assume that \( x^* \) is not a strict relative minimum point; let \( \{ y_k \} \) be a sequence of feasible points converging to \( x^* \) such that \( f(y_k) \leq f(x^*) \), and write each \( y_k \) in the form \( y_k = x^* + \delta_k s_k \) with \( |s_k| = 1, \delta_k > 0 \). We may assume that \( \delta_k \to 0 \) and \( s_k \to s^* \). We have \( 0 \geq \nabla f(x^*)s^* \), and for each \( i = 1, \ldots, m \) we have
\[
\nabla h_i(x^*)s^* = 0.
\]
Also for each active constraint \( g_j \) we have \( g_j(y_k) - g_j(x^*) \leq 0 \), and hence
\[
\nabla g_j(x^*)s^* \leq 0.
\]

If \( \nabla g_j(x^*)s^* = 0 \) for all \( j \in J \), then the proof goes through just as in Section 11.5. If \( \nabla g_j(x^*)s^* < 0 \) for at least one \( j \in J \), then
\[
0 \geq \nabla f(x^*)s^* = -\lambda^T \nabla h(x^*)s^* - \mu^T \nabla g(x^*)s^* > 0,
\]
which is a contradiction. \( \blacksquare \)

We note in particular that if all active inequality constraints have strictly positive corresponding Lagrange multipliers (no degenerate inequalities), then the set \( J \) includes all of the active inequalities. In this case the sufficient condition is that the Lagrangian be positive definite on \( M \), the tangent plane of active constraints.

Sensitivity

The sensitivity result for problems with inequalities is a simple restatement of the result for equalities. In this case, a nondegeneracy assumption is introduced so that the small variations produced in Lagrange multipliers when the constraints are varied will not violate the positivity requirement.

Sensitivity Theorem. Let \( f, g, h \in C^2 \) and consider the family of problems

\[
\begin{align*}
\text{minimize} & \quad f(x) \\
\text{subject to} & \quad h(x) = c \\
& \quad g(x) \leq d.
\end{align*}
\]

Suppose that for \( c = 0, d = 0 \), there is a local solution \( x^* \) that is a regular point and that, together with the associated Lagrange multipliers, \( \lambda, \mu \geq 0 \), satisfies the second-order sufficiency conditions for a strict local minimum. Assume further that no active inequality constraint is degenerate. Then for every \( (c, d) \in E^{m+p} \) in a region containing \( (0, 0) \) there is a solution \( x(c, d) \), depending continuously on \( (c, d) \), such that \( x(0, 0) = x^* \), and such that \( x(c, d) \) is a relative minimum point of (42). Furthermore,

\[
\begin{align*}
\nabla_c f(x(c, d)) \bigg|_{0,0} &= -\lambda^T \\
\nabla_d f(x(c, d)) \bigg|_{0,0} &= -\mu^T.
\end{align*}
\]