CHAPTER 6
CONSTRAINED OPTIMIZATION 1: K-T CONDITIONS

1 Introduction

We now begin our discussion of gradient-based constrained optimization. Recall that in Chapter 3 we looked at gradient-based unconstrained optimization and learned about the necessary and sufficient conditions for an unconstrained optimum, various search directions, conducting a line search, and quasi-Newton methods. We will build on that foundation as we extend the theory to problems with constraints.

2 Necessary Conditions for Constrained Optimum

At an *unconstrained* local optimum, there is no direction in which we can move to improve the objective function. We can state the necessary conditions mathematically as $\nabla f = 0$. At a *constrained* local optimum, there is no *feasible* direction in which we can move to improve the objective. That is, there may be directions from the current point that will improve the objective, but these directions point into infeasible space.

The necessary conditions for a constrained local optimum are called the *Kuhn-Tucker Conditions*, and these conditions play a very important role in constrained optimization theory and algorithm development.

2.1 Problem Form

It will be convenient to cast our optimization problem into one of two particular forms. This is no restriction since any problem can be cast into either of these forms.

$$\begin{aligned} &\text{Max} \quad f(\mathbf{x}) \\ &\text{s.t.:} \\ &\qquad g_i(\mathbf{x}) - b_i \le 0 \quad i = 1, \ \ldots, \ k \\ &\qquad g_i(\mathbf{x}) - b_i = 0 \quad i = k+1, \ \ldots, m \end{aligned}$$

or

$$\begin{aligned} &\text{Min} \quad f(\mathbf{x}) \\ &\text{s.t.:} \\ &\qquad g_i(\mathbf{x}) - b_i \ge 0 \quad i = 1, \ \ldots, \ k \\ &\qquad g_i(\mathbf{x}) - b_i = 0 \quad i = k+1, \ \ldots, m \end{aligned}$$

2.2 Graphical Examples

For the graphical examples below, we will assume we are maximizing with $\le$ constraints.

We have previously considered how we can tell mathematically if some arbitrary vector, $\mathbf{s}$, points downhill. That condition is, $\mathbf{s}^{\mathrm{T}} \nabla f < 0$. We developed this condition by noting that any vector **s** could be resolved into vector components which lie in the tangent plane and along the gradient (or negative gradient) direction.

Now suppose we have the situation shown in Fig. 6.1 below. We are maximizing. We have contours increasing in the direction of the arrow. The gradient vector is shown. What is the set of directions which improves the objective? It is the set for which $\mathbf{s}^{\mathrm{T}}\nabla f > 0$. We show that set as a semi-circle in Fig. 6.1

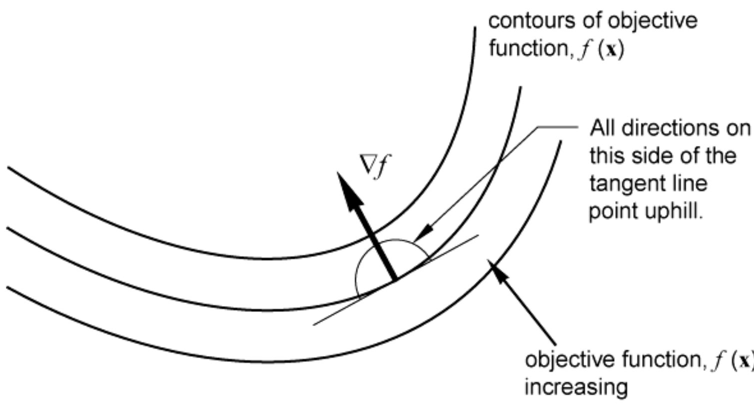

Fig. 6.1 Gradient and set of directions which improves objective function.

Now suppose we add in a less-than inequality constraint, $g(\mathbf{x}) \le 0$. Contours for this constraint are given in Fig. 6.2. The triangular markers indicate the contour for the allowable value and point towards the direction of the feasible space. What is the set of directions which is feasible? It is the set for which $\mathbf{s}^{\mathrm{T}}\nabla f < 0$. That set is shown as a semi-circle in the figure.

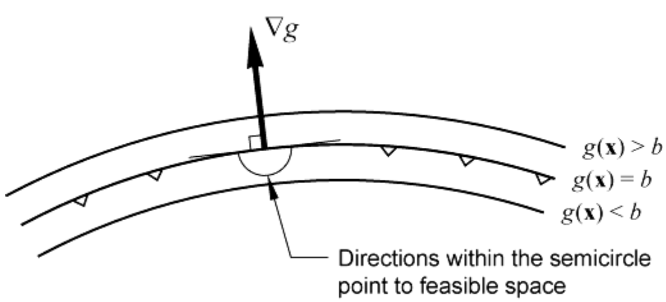

Fig. 6.2 Gradient and set of feasible directions for a constraint.

Now suppose we overlay these two sets of contours on top of each other, as in Fig. 6.3. Where does the optimum lie? By definition, a constrained optimum is a point for which there is no feasible direction which improves the objective. We can see that that condition occurs when the gradient for the objective and gradient for the constraint lie on top of each other. When this happens, the set of directions which improves the objective (dashed semi-circle) does not overlap with the set of feasible directions (solid semi-circle.)

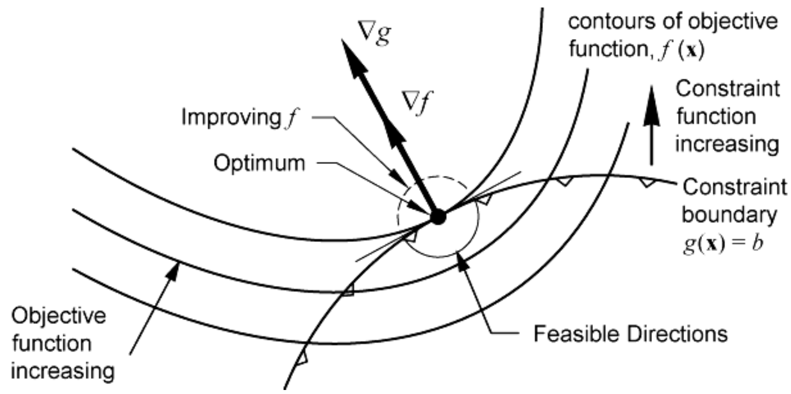

Fig. 6.3 An optimum for one binding constraint occurs when the gradient vectors overlap. When this condition occurs, no feasible point exists which improves the objective.

Mathematically we can write the above condition as

$$\nabla f\left(\mathbf{x}^*\right) = \lambda \nabla g_1\left(\mathbf{x}^*\right) \tag{6.1}$$

where λ is a positive constant.

Now consider a case where there are two binding constraints at the solution, as shown in Fig. 6.4

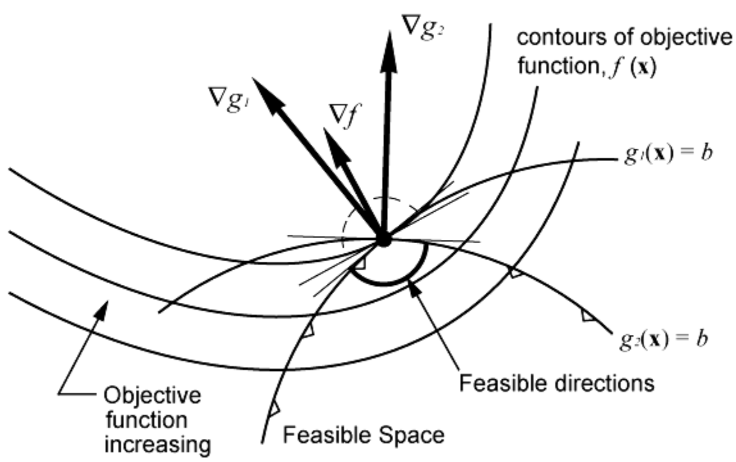

Fig. 6.4 Two binding constraints at an optimum. As long as the objective gradient is within the cone of the constraint gradients, no feasible point exists which improves the objective.

We see that the objective gradient vector is "contained inside" the constraint gradient vectors. If the objective gradient vector is within the constraint gradient vectors, then no direction exists which simultaneously improves the objective and points in the feasible region. We can state this condition mathematically as:

$$\nabla f\left(\mathbf{x}^*\right)=\lambda_1 \nabla g_1\left(\mathbf{x}^*\right)+\lambda_2 \nabla g_2\left(\mathbf{x}^*\right) \tag{6.2}$$

where, as with the single constraint case, λ_1 and λ_2 are positive constants. Having graphically motivated the development of the main mathematical conditions for a constrained optimum, we are now ready to state these conditions.

2.3 The Kuhn-Tucker Conditions

The Kuhn-Tucker conditions are the necessary conditions for a point to be a constrained local optimum, for either of the general problems given below. (The K-T equations also work for an unconstrained optimum, as we will explain later.)

If $\mathbf{x}^*$ is a local max for:

$$\text{Max} \quad f(\mathbf{x}) \tag{6.3}$$

s.t.:

$$g_i(\mathbf{x})-b_i \le 0 \quad i=1,\ \ldots,\ k \tag{6.4}$$

$$g_i(\mathbf{x})-b_i = 0 \quad i=k+1,\ \ldots, m \tag{6.5}$$

Or if $\mathbf{x}^*$ is a local min for:

$$\text{Min} \quad f(\mathbf{x}) \tag{6.6}$$

s.t.:

$$g_i(\mathbf{x})-b_i \ge 0 \quad i=1,\ \ldots,\ k \tag{6.7}$$

$$g_i(\mathbf{x})-b_i = 0 \quad i=k+1,\ \ldots, m \tag{6.8}$$

and if the constraint gradients at the optimum, $\nabla g_i\left(\mathbf{x}^*\right)$, are independent, then there exist $\left(\boldsymbol{\lambda}^*\right)^{\mathrm{T}}=\left[\lambda_1 \ldots \lambda_m\right]$, called *Lagrange multipliers*, such that $\mathbf{x}^*$ and $\boldsymbol{\lambda}^*$ satisfy the following system of equations,

$$g_i\left(\mathbf{x}^*\right)-b_i \text{ is feasible} \quad i=1,\ldots,m \tag{6.9}$$

$$\nabla f\left(\mathbf{x}^*\right)-\sum_{i=1}^{m} \lambda_i^* \nabla g_i\left(\mathbf{x}^*\right)=\mathbf{0} \tag{6.10}$$

$$\lambda_i^*\left[g_i\left(\mathbf{x}^*\right)-b_i\right]=0 \quad i=1,\ldots,k \tag{6.11}$$

$$\lambda_i^* \ge 0 \quad i=1,\ldots,k \tag{6.12}$$

$$\begin{gathered}\lambda_i^* \text{ unrestricted for } i=k+1,\ldots,m \\ \text{or } -\infty<\lambda_i<\infty\end{gathered} \tag{6.13}$$

Note in the above equations, $i = 1, \ldots, k$ indicates inequality constraints, $i = k+1, \ldots, m$ indicates equality constraints, and $i = 1, \ldots, m$ indicates all constraints.

Just as with the necessary conditions for an unconstrained optimum, the K-T conditions are necessary but not sufficient conditions for a constrained optimum.

We will now explain each of these conditions.

Equation (6.9) requires that a constrained optimum be feasible with respect to all constraints.

Equation (6.10) requires the objective function gradient to be a linear combination of the constraint gradients. This insures there is no direction that will simultaneously improve the objective and satisfy the constraints.

Equation (6.11) enforces a condition known as *complementary slackness.* Notice that this condition is for the inequality constraints only. This condition states that either an inequality constraint is binding, or the associated Lagrange multiplier is zero. Essentially this means that nonbinding inequality constraints drop out of the problem.

Equation (6.12) states that the Lagrange multipliers for the inequality constraints must be positive.

Equation (6.13) states that the Lagrange multipliers for the equality constraints can be either positive or negative.

Note that (6.10) above, which is given in vector form, represents a system of *n* equations. We can rewrite (6.10) as:

$$\begin{array}{cccccc} \dfrac{\partial f}{\partial x_1} & -\lambda_1 \dfrac{\partial g_1}{\partial x_1} & -\lambda_2 \dfrac{\partial g_2}{\partial x_1} & - \cdots - & \lambda_m \dfrac{\partial g_m}{\partial x_1} & = 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \dfrac{\partial f}{\partial x_n} & -\lambda_1 \dfrac{\partial g_1}{\partial x_n} & -\lambda_2 \dfrac{\partial g_2}{\partial x_n} & - \cdots - & \lambda_m \dfrac{\partial g_m}{\partial x_n} & = 0 \end{array} \tag{6.14}$$

We note there is a Lagrange multiplier, λ, for every constraint. Recall, however, that if the constraint is not binding then its Lagrange multiplier is zero, from (6.11).

Taken together, the K-T conditions represent $m+n$ equations in $m+n$ unknowns. The equations are the *n* equations given by (6.14) (or (6.10)) and the *m* constraints ((6.9)). The unknowns are the *n* elements of the vector $\mathbf{x}$ and the *m* elements of the vector $\boldsymbol{\lambda}$

2.4 Examples of the K-T Conditions

2.4.1 Example 1: An Equality Constrained Problem

Using the K-T equations, find the optimum to the problem,

$$\text{Min} \quad f(\mathbf{x}) = 2x_1^2 + 4x_2^2$$
$$\text{s.t.} \quad g_1: \; 3x_1 + 2x_2 = 12$$

A picture of this problem is given below:

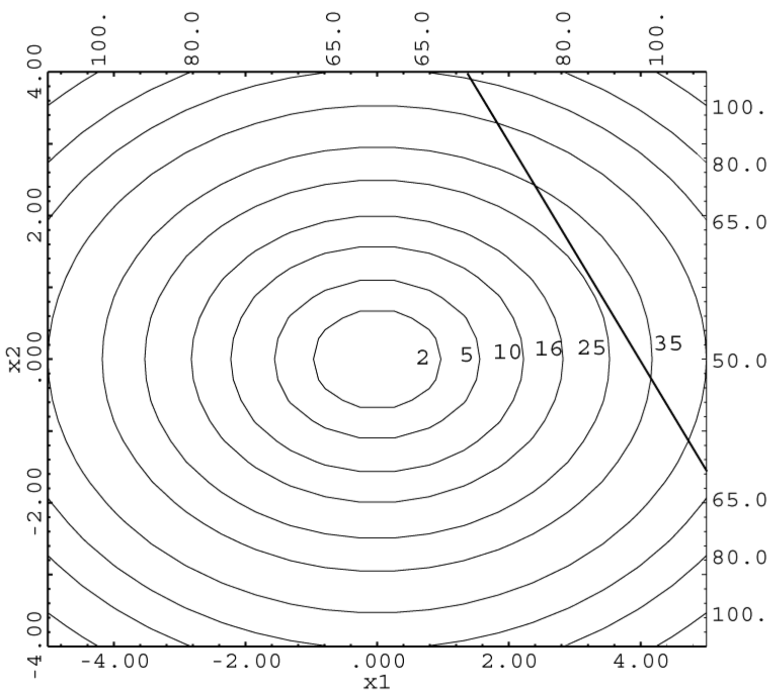

Fig. 6.5 Contours of functions for Example 1.

Since the constraint is an equality constraint, we know it is binding, so the Lagrange multiplier will be non-zero. With two variables and one constraint, the K-T equations represent three equations in three unknowns.

The K-T conditions can be written:

$$\frac{\partial f}{\partial x_1} - \lambda \frac{\partial g_1}{\partial x_1} = 0$$
$$\frac{\partial f}{\partial x_2} - \lambda \frac{\partial g_1}{\partial x_2} = 0$$
$$g_1(\mathbf{x}) - b_1 = 0$$

evaluating these expressions:

$$4x_1 - \lambda(3) = 0$$
$$8x_2 - \lambda(2) = 0$$
$$3x_1 + 2x_2 - 12 = 0$$

which we can write in matrix form as:

$$\begin{bmatrix} 4 & 0 & -3 \\ 0 & 8 & -2 \\ 3 & 2 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \lambda \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 12 \end{bmatrix}$$

The solution to this system of equations is $x_1 = 3.2727,\ x_2 = 1.0909,\ \lambda = 4.3636$. The value of the objective at this point is $f = 26.18$. This optimum occurs where the gradient vectors of the constraint and objective overlap, just as indicated by the graphical discussion. We should verify to make sure this is a constrained min and not a constrained max, since the K-T equations apply at both.

Because the objective function for this problem is quadratic and the constraint is linear, the K-T equations are linear and thus easy to solve. We call a problem with a quadratic objective and linear constraints a *quadratic programming problem* for this reason. Usually the K-T equations are not linear. However, the SQP algorithm attempts to solve these equations by solving a sequence of quadratic program approximations to the real program—thus the name of "Sequential Quadratic Programming."

2.4.2 Example 2: An Inequality Constrained Problem

In general it is more difficult to use the K-T conditions to solve for the optimum of an inequality constrained problem (than for a problem with equality constraints only) because we don't know beforehand which constraints are binding at the optimum. Thus we often use the K-T conditions to *verify* that a point we have reached is a candidate optimal solution. Given a point, it is easy to check which constraints are binding.

Verify that the point $\mathbf{x}^{\mathrm{T}} = \begin{bmatrix} 0.7059 & 2.8235 \end{bmatrix}$ is an optimum to the problem:

$$\begin{aligned} &\text{Min} && f(\mathbf{x}) = x_1^2 + x_2^2 \\ &\text{s.t.} && g:\ \ x_1 + 4x_2 \ge 12 \end{aligned}$$

Step 1: Put problem in proper form:

$$\begin{aligned} &\text{Min} && f(\mathbf{x}) = x_1^2 + x_2^2 \\ &\text{s.t.} && g = x_1 + 4x_2 - 12 \ge 0 \end{aligned}$$

Step 2: See which constraints are binding:

$$0.7059+4(2.8235)-12=\text{-}0.0001 \quad \Rightarrow \quad \lambda \neq 0$$

Since this constraint is binding, the associated Lagrange multiplier is solved for. (If it were not binding, the Lagrange multiplier would be zero, from complementary slackness.)

Step 3: Write out the Lagrange multiplier equations represented by (6.10):

$$\frac{\partial f}{\partial x_1}-\lambda\frac{\partial g_1}{\partial x_1}=2x_1-\lambda(1)=0$$
$$\frac{\partial f}{\partial x_2}-\lambda\frac{\partial g_1}{\partial x_2}=2x_2-\lambda(4)=0$$

Step 4: Substitute in the given point:

$$2(0.7059)=\lambda \tag{6.15}$$
$$2(2.8235)=4\lambda \tag{6.16}$$

From (6.15), $\lambda=1.4118$
From (6.16), $\lambda=1.4118$

Since these λ's are consistent and positive, the above point satisfies the K-T equations and is a candidate optimal solution.

2.4.3 Example 3: Another Inequality Constrained Problem

Given the problem:

$$\begin{array}{ll} \text{Min} & f(\mathbf{x})=x_1^2+x_2 \\ \text{s.t.} & g_1(\mathbf{x})=x_1^2+x_2^2-9\le 0 \\ & g_2(x)=x_1+x_2-1\le 0 \end{array}$$

See if $\mathbf{x}^*=\begin{bmatrix}0 & -3\end{bmatrix}^{\mathrm{T}}$ satisfies the K–T conditions.

Graphically the problem looks like,

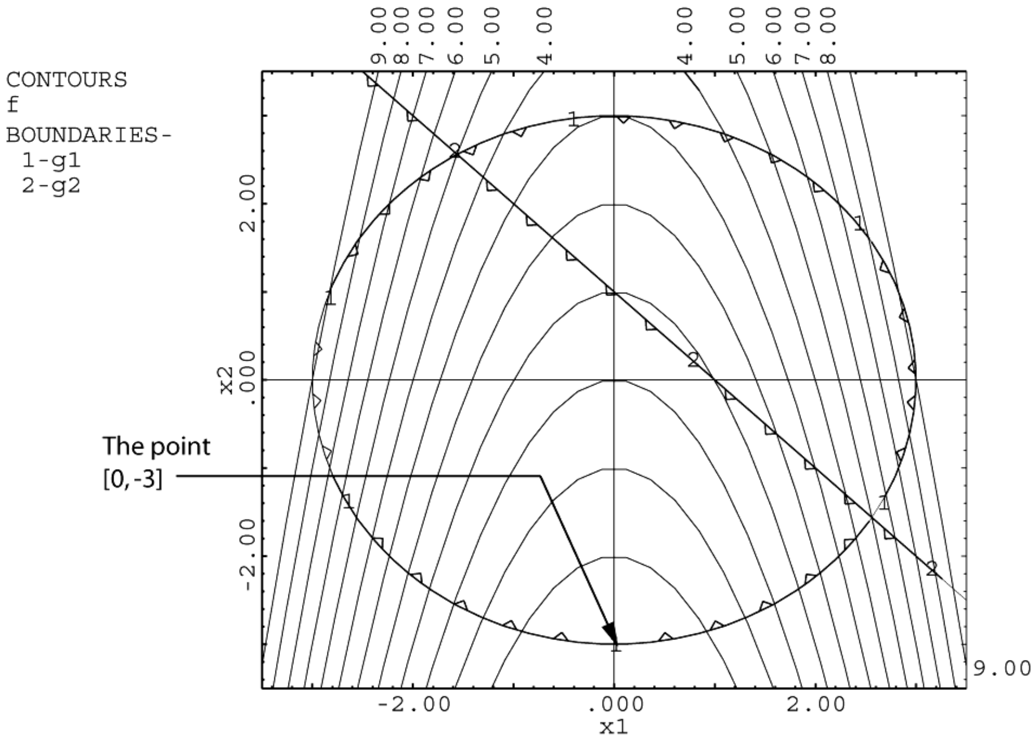

Fig. 6.6. Contour plot and proposed point for Example 3.

At the proposed point constraint g_1 is binding; g_2 is not.

Step 1: Change problem to be in the form of (6.3-6.5):

$$\begin{aligned} &\text{Max} && f(\mathbf{x}) = -x_1^2 - x_2 \\ &\text{s.t.} && g_1(\mathbf{x}): \quad x_1^2 + x_2^2 - 9 \le 0 \\ &&& g_2(\mathbf{x}): \quad x_1 + x_2 - 1 \le 0 \end{aligned}$$

Step 2: See which constraints are binding:

In this case we can check constraints graphically. Because $g_2(\mathbf{x})$ is not binding, $\lambda_2 = 0$ from (6.11). However, λ_1 is solved for since $g_1(\mathbf{x})$ is binding.

Step 3: Write out the Lagrange multiplier equations represented by (6.10):

$$\frac{\partial f}{\partial x_1} - \lambda_1 \frac{\partial g_1}{\partial x_1} - \lambda_2 \frac{\partial g_2}{\partial x_1} = -2x_1 - \lambda_1(2x_1) = 0 \tag{6.17}$$

$$\frac{\partial f}{\partial x_2} - \lambda_1 \frac{\partial g_1}{\partial x_2} - \lambda_2 \frac{\partial g_2}{\partial x_2} = -1 - \lambda_1(2x_2) = 0 \tag{6.18}$$

Step 4: Substitute in the given point:

At $\mathbf{x}^{\mathrm{T}} = \begin{bmatrix} 0 & -3 \end{bmatrix}$, (6.17) vanishes; from (6.18):

$$-\lambda_1(2)(-3) = 1 \quad \Rightarrow \quad \lambda_1 = \frac{1}{6}$$

so a valid set of λ's, i.e., $\left(\lambda^*\right)^{\mathrm{T}} = \begin{bmatrix} \frac{1}{6} & 0 \end{bmatrix}$ has been found and the K–T conditions are satisfied. This is therefore a candidate optimal solution.

2.4.4 Example 4: Another Point for the Same Problem

Check to see if $\left(\mathbf{x}^*\right)^{\mathrm{T}} = \begin{bmatrix} 1 & 0 \end{bmatrix}$ satisfies the K–T conditions for the problem given in Example 3 above. This point is shown in Fig. 6.7

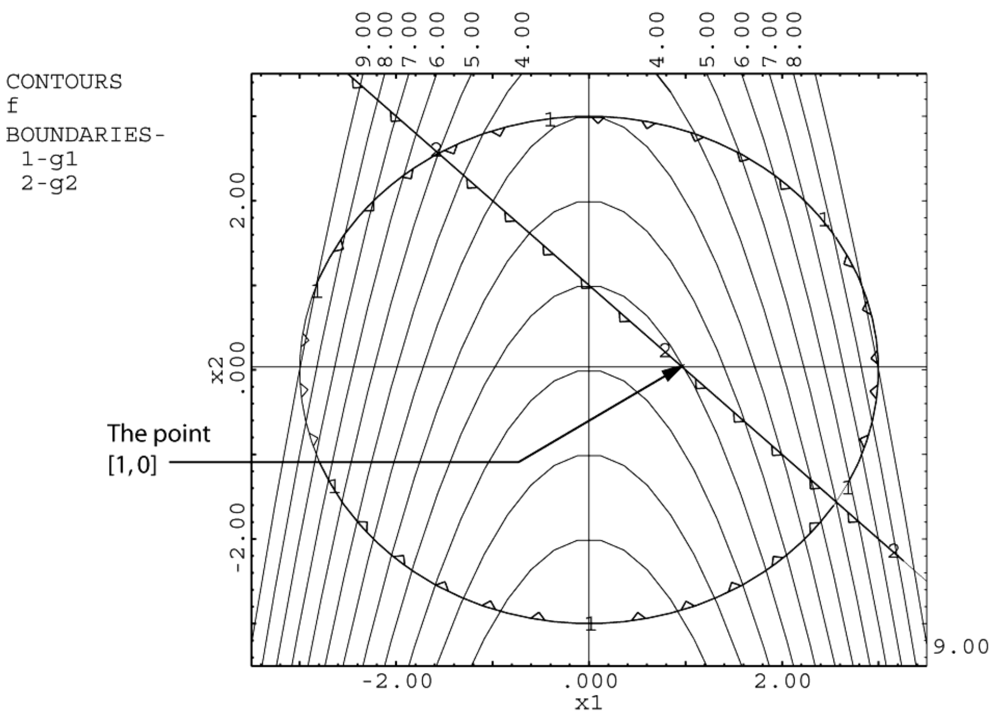

Fig. 6.7 Contour plot and proposed point for Example 4.

Step 1: Change problem to be in the form of (6.3-6.5):

$$\begin{aligned} &\text{Max} && f(\mathbf{x}) = -x_1^2 - x_2 \\ &\text{s.t.} && g_1(\mathbf{x}): \quad x_1^2 + x_2^2 - 9 \le 0 \\ & && g_2(\mathbf{x}): \quad x_1 + x_2 - 1 \le 0 \end{aligned}$$

Step 2: See which constraints are binding:

From Fig. 6.7 we can see that $g_1(\mathbf{x})$ is not binding and therefore $\lambda_1 = 0$; $g_2(\mathbf{x})$ is binding so $\lambda_2 \neq 0$

Step 3: Write out the Lagrange multiplier equations represented by (6.10):

$$\frac{\partial f}{\partial x_1} - \lambda_1 \frac{\partial g_1}{\partial x_1} - \lambda_2 \frac{\partial g_2}{\partial x_1} = -2x_1 - \lambda_2(1) = 0 \tag{6.19}$$

$$\frac{\partial f}{\partial x_2} - \lambda_1 \frac{\partial g_1}{\partial x_2} - \lambda_2 \frac{\partial g_2}{\partial x_2} = -1 - \lambda_2(1) = 0 \tag{6.20}$$

Step 4: Substitute in the given point:

Substituting $\mathbf{x}^{\mathrm{T}} = \begin{bmatrix} 1 & 0 \end{bmatrix}$

$\lambda_2 = -2$ from (6.19)

$\lambda_2 = -1$ from (6.20)

Since we cannot find a consistent set of λ's, and the λ's are negative as well (either condition being enough to disqualify the point), this point does not satisfy the Kuhn-Tucker conditions and *cannot* be a constrained optimum.

Question: In Examples 3 and 4 we have looked at two points—a constrained min, and point which is not an optimum. Are there any other points which would satisfy the K-T conditions for this problem? Where would a constrained max be found? Would the K-T conditions apply there?

2.5 Unconstrained Problems

We mentioned the K-T conditions also apply to unconstrained problems. This is fortunate since a constrained optimization problem does not have to have a constrained solution. The optimum might be an unconstrained optimum in the interior of the constraints.

If we have an unconstrained optimum to a constrained problem, what happens to the K-T conditions? In this case none of the constraints are binding so *all* of the Lagrange multipliers are zero, from (6.11), so (6.10) becomes,

$$\nabla f\left(\mathbf{x}^*\right) - \left[\cancelto{=0}{\sum_{i=1}^{m} \lambda_i^* \nabla g_i\left(\mathbf{x}^*\right)}\right] = \nabla f\left(\mathbf{x}^*\right) = 0$$

Thus we see that the K-T equations simplify to the necessary conditions for an unconstrained optimum when no constraints are binding.

3 The Lagrangian Function

3.1 Definition

It will be convenient for us to define the *Lagrangian Function*:

$$\mathrm{L}(\mathbf{x},\boldsymbol{\lambda}) = f(\mathbf{x}) - \sum_{i=1}^{m} \lambda_i \left[g_i(\mathbf{x}) - b_i \right] \tag{6.21}$$

Note that the Lagrangian function is a function of both **x** and **λ**. Thus the gradient of the Lagrangian function is made up of partials with respect to x and λ:

$$\left[\nabla \mathrm{L}(\mathbf{x},\boldsymbol{\lambda})\right]^{\mathrm{T}} = \left[\frac{\partial \mathrm{L}}{\partial x_1} \quad \cdots \quad \frac{\partial \mathrm{L}}{\partial x_n} \quad \frac{\partial \mathrm{L}}{\partial \lambda_1} \quad \cdots \quad \frac{\partial \mathrm{L}}{\partial \lambda_m} \right] \tag{6.22}$$

We will evaluate some of these partials to become familiar with them,

The partial $\frac{\partial \mathrm{L}}{\partial x_1}$ is: $$\frac{\partial \mathrm{L}}{\partial x_1} = \frac{\partial f}{\partial x_1} - \sum_{i=1}^{m} \lambda_i \frac{\partial g_i}{\partial x_1}$$

Similarly, $\frac{\partial \mathrm{L}}{\partial x_2}$ is given by, $$\frac{\partial \mathrm{L}}{\partial x_2} = \frac{\partial f}{\partial x_2} - \sum_{i=1}^{m} \lambda_i \frac{\partial g_i}{\partial x_2}$$

The partial $\frac{\partial \mathrm{L}}{\partial \lambda_1}$ is: $$\frac{\partial \mathrm{L}}{\partial \lambda_1} = -\left[g_1 - b_1 \right]$$

It is convenient, given these results, to split the gradient vector of the Lagrangian function into two parts: the vector containing the partial derivatives with respect to *x*, written $\nabla_{\mathrm{x}} \mathrm{L}$, and the vector containing the partials with respect to λ, written $\nabla_{\lambda} \mathrm{L}$.

The gradient of the Lagrangian function with respect to *x* can be written in vector form as:

$$\nabla_x \mathrm{L} = \nabla f(\mathbf{x}) - \sum_{i=1}^{m} \lambda_i \nabla g_i(\mathbf{x}) \tag{6.23}$$

so that we could replace (6.10) by $\nabla_{\mathrm{x}} \mathrm{L} = \mathbf{0}$ if we wished.

The gradient of the Lagrangian function with respect to λ is:

$$\nabla_\lambda \mathrm{L} = -\begin{bmatrix} g_1(\mathbf{x}) - b_1 \\ g_2(\mathbf{x}) - b_2 \\ \vdots \\ g_m(\mathbf{x}) - b_m \end{bmatrix} \tag{6.24}$$

3.2 The Lagrangian Function and Optimality

For a problem with equality constraints only, we can compactly state the K-T conditions as,

$$\nabla \mathrm{L} = \begin{bmatrix} \nabla_x \mathrm{L} \\ \nabla_\lambda \mathrm{L} \end{bmatrix} = \mathbf{0} \tag{6.25}$$

For a problem with inequality constraints as well, the main condition of the K-T equations, (6.10), can be stated as,

$$\nabla_x \mathrm{L} = \mathbf{0} \tag{6.26}$$

Thus we can consider that at an optimum, there exist **λ*** and **x*** such that **x*** is a stationary point of the Lagrangian function.

The Lagrangian function provides information about how the objective and binding constraints together affect an optimum. Suppose we are at a constrained optimum. If we were to change the objective function, this would clearly have an effect on the solution. Likewise, if we were to change the binding constraints (perhaps by changing the right hand sides), these changes would also affect the value of the solution. The Lagrangian function tells how these changes trade-off against each other,

$$\mathrm{L}(\mathbf{x},\boldsymbol{\lambda}) = f(\mathbf{x}) - \sum_{i=1}^{m} \lambda_i \left[g_i(\mathbf{x}) - b_i \right]$$

The Lagrange multipliers serve as "weighting factors" between the individual constraints and the objective. Appropriately, the multipliers have units of (objective function/constraint function). Thus if our objective had units of pounds, and constraint *i* had units of inches, Lagrange multiplier *i* would have units of pounds per inch.

3.3 Interpreting Values of Lagrange Multipliers

Thus far we have solved for Lagrange multipliers, but we have not attached any significance to their values. We will use the Lagrangian function to help us interpret what their values mean.

We will start with the Lagrangian function at an optimum:

$$L\left(\mathbf{x}^*,\boldsymbol{\lambda}^*\right) = f\left(\mathbf{x}^*\right) - \sum_{i=1}^{m} \lambda_i^* \left[g_i\left(\mathbf{x}^*\right) - b_i \right] \tag{6.27}$$

Suppose now we consider the right-hand side of constraint *i*, b_i, to be a variable. How does the optimal solution change as we change b_i? To answer this question, we need to find $\frac{df^*}{db_i}$.

That is, we need to find how the *optimal* value of the objective changes as we change the right hand side, b_i. The Lagrangian function, which relates how constraints and the objective interact at the optimum, can give us this information.

We will be concerned only with small perturbations at the optimum. This allows us to ignore *nonbinding* inequality constraints, which will be treated as if they were not there. Thus instead of *m* constraints, we will have *m** constraints, which is the set of equality and binding inequality constraints.

At an optimum, the value of L becomes the same as the value of *f*. This is because all of the terms in braces go to zero,

$$L\left(\mathbf{x}^*,\boldsymbol{\lambda}^*\right) = f\left(\mathbf{x}^*\right) - \left[\overset{=0}{\cancel{\sum_{i=1}^{m^*} \lambda_i^* \left[g_i\left(\mathbf{x}^*\right) - b_i \right]}} \right] = f\left(\mathbf{x}^*\right)$$

since all constraints in our set m* are binding. At the optimum therefore, $\frac{df^*}{db_i} = \frac{dL^*}{db_i}$.

As we change b_i, we would expect $\mathbf{x}^*$ and $\boldsymbol{\lambda}^*$ to change, so we need to consider $\mathbf{x}$ and $\boldsymbol{\lambda}$ themselves to be functions of b_i, i.e., $\mathbf{x}\left(b_i\right)$, $\boldsymbol{\lambda}\left(b_i\right)$. Then by the chain rule:

$$\frac{df^*}{db_i} = \frac{dL^*}{db_i} = \frac{\partial \mathbf{x}^{\mathrm{T}}}{\partial b_i} \nabla_x L + \frac{\partial \boldsymbol{\lambda}^{\mathrm{T}}}{\partial b_i} \nabla_\lambda L + \frac{\partial L}{\partial b_i} \tag{6.28}$$

At the optimum $\nabla_x L = 0$ and $\nabla_\lambda L = 0$, leaving only $\frac{\partial L}{\partial b_i}$.

From the Lagrangian function, $\frac{\partial L}{\partial b_i} = \lambda_i$,

Thus we have the result,

$$\frac{df^*}{db_i} = \lambda_i^* \tag{6.29}$$

The Lagrange multipliers provide us with *sensitivity information* about the optimal solution. They tell use how the optimal objective value would change if we changed the right-hand side of the constraints. This can be very useful information, telling us, for example, how we could expect the optimal weight of a truss to change if we relaxed the right-hand side of a binding stress constraint (i.e. increased the allowable value of the stress constraint).

<u>Caution</u>: The sensitivity information provided by the Lagrange multipliers is valid only for small changes about the optimum (since for nonlinear functions, derivatives change as we move from point to point) and assumes *that the same constraints* are binding at the perturbed optimum.

3.3.1 <u>Example 5: Interpreting the Value of the Lagrange Multipliers</u>

In Example 1, Section 2.4.1, we solved the following problem

$$\begin{array}{ll} \text{Min} & f(\mathbf{x}) = 2x_1^2 + 4x_2^2 \\ \text{s.t.} & g_1: \; 3x_1 + 2x_2 = 12 \end{array}$$

We found the optimal solution to be:
$x_1 = 3.2727,\ x_2 = 1.0909,\ \lambda = 4.3636$, at which point $f^* = 26.18$.

What would be the expected change in the objective be if we increased the right-hand side of the constraint from 12 to 13? From (6.29),

$$\Delta f^* \approx \lambda_i^* \Delta b_i$$

For $\Delta b = 1$, the change in the objective should be approximately 4.36.

If we change the right hand side and re-optimize the problem, the new optimum is, $x_1 = 3.5454,\ x_2 = 1.1818,\ \lambda = 4.7272$, at which point $f^* = 30.72$. The actual change in the objective is 4.54. (Indeed, it is the average of the two $\lambda's$.)

Thus, without optimizing, the Lagrange multiplier provides an estimate of how much the objective would change per unit change in the constraint right hand side. This helps us evaluate how sensitive the optimum is to changes.

Sometimes the objective represents profit, and the right hand side of a constraint is viewed as a *resource* which can be purchased. The value of the Lagrange multiplier is a breakpoint between realizing a net increase in profit or a net loss. If, for a binding constraint, we can purchase more right hand side for less than the Lagrange multiplier, net profit will be positive. If not, the cost of the resource will outweigh the increase in profit.

3.4 <u>Necessary and Sufficient Conditions</u>

The K-T Conditions we have presented in previous sections are necessary conditions for a constrained optimum. That is, for a point to be a candidate optimal solution, it must be possible to find values of $\boldsymbol{\lambda}$ that satisfy (6.9)-(6.13). If we cannot find such $\boldsymbol{\lambda}$, then the candidate point *cannot be* a constrained optimal solution.

Note, however, that as part of the KT conditions we require the constraint gradients, $\nabla g_i\left(x^*\right)$, to be independent at the optimal solution; otherwise it is possible, although unlikely, that we could have a point be a constrained optimum and not satisfy the KT conditions.

If a point satisfies the KT conditions, then it is a *candidate* optimal solution. As we have seen, the necessary conditions can hold at a constrained max, constrained min, or a point that is neither. An example showing how this might occur is given below:

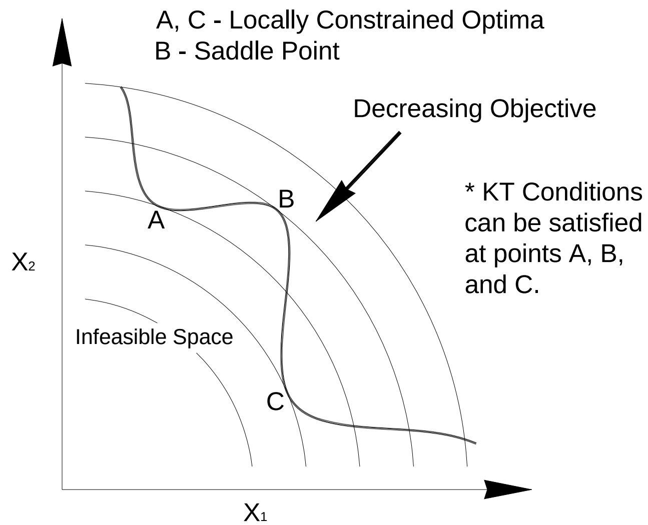

Fig. 6.8. Points where the K-T equations would be satisfied.

For an unconstrained optimum we saw that *sufficient* conditions for a minimum were that $\nabla f = 0$ and, the Hessian, $\nabla^2 f(\mathbf{x})$, is positive definite.

Likewise, for a constrained optimum, sufficient conditions for a point to be a *constrained minimum* are the K-T equations are satisfied (6-9-6.13) and the Hessian of the Lagrangian function with respect to x, $\nabla_x^2 \mathrm{L}\left(\mathbf{x}^*, \boldsymbol{\lambda}^*\right)$, is positive definite, where,

$$\nabla_x^2 \mathrm{L}\left(\mathbf{x}^*, \boldsymbol{\lambda}^*\right) = \nabla^2 f\left(\mathbf{x}^*\right) - \sum_{i=1}^{m} \lambda_i^* \nabla^2 g_i\left(\mathbf{x}^*\right) \tag{6.30}$$

Some further discussion is needed, however. If we write the condition of positive definiteness as,

$$\mathbf{y}^T \nabla_x^2 \mathrm{L}\left(\mathbf{x}^*, \boldsymbol{\lambda}^*\right)\mathbf{y} > 0 \tag{6.31}$$

The vectors $\mathbf{y}$ must satisfy,

$$\mathbf{J}(\mathbf{x}^*)\mathbf{y} > 0 \tag{6.32}$$

Where $\mathbf{J}(\mathbf{x}^*)$ is the Jacobian matrix of the constraints (matrix whose rows are the gradients of the constraints) active at $\mathbf{x}^*$. These vectors comprise a tangent plane and are orthogonal to the gradients of the active constraints. For more information about the sufficient conditions, see Luenberger, (1984), Fletcher (1987) or Edgar et al. (2001).

4 References

Fletcher, R., *Practical Methods of Optimization 2nd Ed.*, Wiley, 1987.
Luenberger, D. G., *Linear and Nonlinear Programming, 2nd Ed,* Addison Wesley, 1984
Edgar, T. F., Himmelblau, D. M., Lasdon, L. S., *Optimization of Chemical Processes, 2nd Ed.,* McGraw-Hill, 2001.