Chapter 21 Problems with Inequality Constraints

An Introduction to Optimization
Spring, 2014

Wei-Ta Chu

Karush-Kuhn-Tucker Condition

- Consider the following problem:

$$\begin{aligned} \text{minimize } & f(\boldsymbol{x}) \\ \text{subject to } & \boldsymbol{h}(\boldsymbol{x}) = \boldsymbol{0} \\ & \boldsymbol{g}(\boldsymbol{x}) \leq \boldsymbol{0} \end{aligned}$$

  where $f : \mathbb{R}^n \to \mathbb{R}$, $\boldsymbol{h} : \mathbb{R}^n \to \mathbb{R}^m$, $m \leq n$, and $\boldsymbol{g} : \mathbb{R}^n \to \mathbb{R}^p$.

- Definition 21.1. An inequality constraint $g_j(\boldsymbol{x}) \leq 0$ is said to be *active* at $\boldsymbol{x}^*$ if $g_j(\boldsymbol{x}^*) = 0$. It is *inactive* at $\boldsymbol{x}^*$ if $g_j(\boldsymbol{x}^*) < 0$

- Definition 21.2. Let $\boldsymbol{x}^*$ satisfy $\boldsymbol{h}(\boldsymbol{x}^*) = \boldsymbol{0}$, $\boldsymbol{g}(\boldsymbol{x}^*) \leq \boldsymbol{0}$, and let $J(\boldsymbol{x}^*)$ be the index set of active inequality constraints

$$J(\boldsymbol{x}^*) \triangleq \{j : g_j(\boldsymbol{x}^*) = 0\}$$

  Then, we say that $\boldsymbol{x}^*$ is a *regular point* if the vectors $\triangledown h_i(\boldsymbol{x}^*), \triangledown g_j(\boldsymbol{x}^*), 1 \leq i \leq m, j \in J(\boldsymbol{x}^*)$ are linearly independent.

Karush-Kuhn-Tucker Condition

- We now prove a first-order necessary condition for a point to be a local minimizer. We call this condition the Karush-Kuhn-Tucker (KKT) condition (or Kuhn-Tucker condition)
- Theorem 21.1. **Karush-Kuhn-Tucker Theorem**. Let $f, \boldsymbol{h}, \boldsymbol{g} \in \mathcal{C}^1$. Let $\boldsymbol{x}^*$ be a regular point and a local minimizer for the problem of minimizing f subject to $\boldsymbol{h}(\boldsymbol{x}) = \mathbf{0}, \boldsymbol{g}(\boldsymbol{x}) \leq \mathbf{0}$ Then, there exists $\boldsymbol{\lambda}^* \in \mathbb{R}^m$ and $\boldsymbol{\mu}^* \in \mathbb{R}^p$ such that:
  - 1. $\boldsymbol{\mu}^* \geq \mathbf{0}$
  - 2. $Df(\boldsymbol{x}^*) + \boldsymbol{\lambda}^{*T} D\boldsymbol{h}(\boldsymbol{x}^*) + \boldsymbol{\mu}^{*T} D\boldsymbol{g}(\boldsymbol{x}^*) = \mathbf{0}^T$
  - 3. $\boldsymbol{\mu}^{*T} \boldsymbol{g}(\boldsymbol{x}^*) = 0$

Karush-Kuhn-Tucker Condition

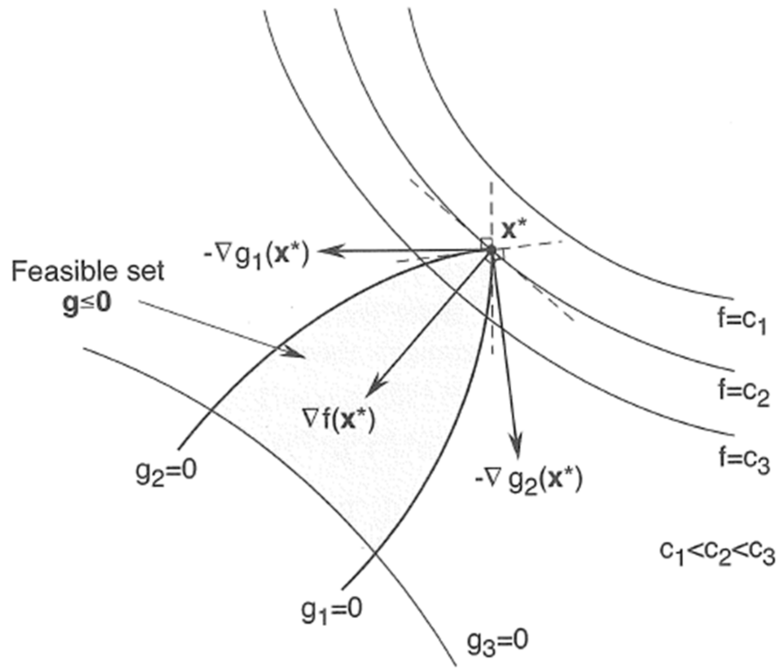
- In Theorem 21.1, we refer to $\boldsymbol{\lambda}^*$ as the Lagrange multiplier vector and $\boldsymbol{\mu}^*$ as the Karush-Kuhn-Tucker (KKT) multiplier vector. We refer to their components as Lagrange multipliers and KKT multipliers, respectively.
- Observe that $\mu_j^* \geq 0$ (by condition 1) and $g_j(\boldsymbol{x}^*) \leq 0$. Therefore, the condition

$$\boldsymbol{u}^{*T}\boldsymbol{g}(\boldsymbol{x}^*) = u_1^* g_1(\boldsymbol{x}^*) + \cdots + u_p^* g_p(\boldsymbol{x}^*) = 0$$

implies that if $g_j(\boldsymbol{x}^*) < 0$, then $\mu_j^* = 0$; that is, for all $j \notin J(\boldsymbol{x}^*)$ we have $\mu_j^* = 0$. In other words, the KKT multipliers μ_j^* corresponding to inactive constraints are zero. The other KKT multipliers, μ_j^*, $j \in J(\boldsymbol{x}^*)$, are nonnegative; they may or may not be equal to zero.

Example 21.1

- A graphical illustration of the KKT theorem is given in Figure 21.1. In this two-dimensional example, we have only inequality constraints $g_j(\boldsymbol{x}^*) \leq 0$, $j = 1, 2, 3$. Note that the point $\boldsymbol{x}^*$ in the figure is indeed a minimizer.
- Figure 21.1



Example 21.1

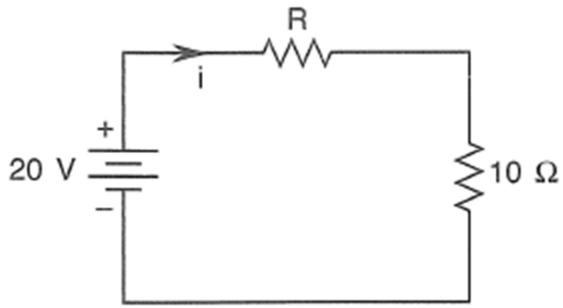
- The constraint $g_3(\boldsymbol{x}) \leq 0$ is inactive: $g_3(\boldsymbol{x}^*) < 0$; hence $\mu_3^* = 0$
By the KKT theorem, we have
$$\nabla f(\boldsymbol{x}^*) + \mu_1^* \nabla g_1(\boldsymbol{x}^*) + \mu_2^* \nabla g_2(\boldsymbol{x}^*) = \mathbf{0}$$
or, equivalently, $\nabla f(\boldsymbol{x}^*) = -\mu_1^* \nabla g_1(\boldsymbol{x}^*) - \mu_2^* \nabla g_2(\boldsymbol{x}^*)$
where $\mu_1^* > 0, \mu_2^* > 0$
- It is easy to interpret the KKT condition graphically for this example. Specifically, we can see from Figure 21.1 that $\nabla f(\boldsymbol{x}^*)$ must be a linear combination of the vectors $-\nabla g_1(\boldsymbol{x}^*)$ and $-\nabla g_2(\boldsymbol{x}^*)$ with positive coefficients. This is reflected exactly in the equation above, where the coefficients μ_1^*, μ_2^* are the KKT multipliers.

Karush-Kuhn-Tucker Condition

- We apply the KKT condition in the same way that we apply any necessary condition. Specifically, we search for points satisfying the KKT condition and treat these points as candidate minimizers. To summarize, the KKT condition consists of five parts (three equations and two inequalities):
  - 1. $\boldsymbol{\mu}^* \geq \mathbf{0}$
  - 2. $Df(\boldsymbol{x}^*) + \boldsymbol{\lambda}^{*T} D\boldsymbol{h}(\boldsymbol{x}^*) + \boldsymbol{\mu}^{*T} D\boldsymbol{g}(\boldsymbol{x}^*) = \mathbf{0}^T$
  - 3. $\boldsymbol{\mu}^{*T} \boldsymbol{g}(\boldsymbol{x}^*) = 0$
  - 4. $\boldsymbol{h}(\boldsymbol{x}^*) = \mathbf{0}$
  - 5. $\boldsymbol{g}(\boldsymbol{x}^*) \leq \mathbf{0}$

Example 21.2

- Consider the circuit in Figure 21.2. Formulate and solve the KKT condition for the following problems.
  - 1. Find the value of the resistor $R \geq 0$ such that the power absorbed by this resistor is maximized.
  - 2. Find the value of the resistor $R \geq 0$ such that the power delivered to the $10-\Omega$ resistor is maximized.
- Figure 21.2



Example 21.2

- The power absorbed by the resistor R is $p = i^2 R$, where $i = \frac{20}{10+R}$. The optimization problem can be represented as

$$\begin{aligned} &\text{minimize} \quad -\frac{400R}{(10+R)^2} \\ &\text{subject to} \quad -R \leq 0 \end{aligned}$$

- The derivative of the objective function is

$$-\frac{400(10+R)^2 - 800R(10+R)}{(10+R)^4} = -\frac{400(10-R)}{(10+R)^3}$$

Thus, the KKT condition is

$$-\frac{400(10-R)}{(10+R)^3} - \mu = 0$$

$$\mu \geq 0, \quad \mu R = 0, \quad -R \leq 0$$

Example 21.2

- We consider two cases. In the first case, suppose that $\mu > 0$ Then, $R = 0$. But this contradicts the first condition above. Now suppose that $\mu = 0$. Then, by the first condition, we have $R = 10$. Therefore, the only solution to the KKT condition is $R = 10, \mu = 0$

Example 21.2

- The power absorbed by the $10-\Omega$ resistor is $p = i^2 10$, where $i = 20/(10+R)$. The optimization problem can be represented as

$$\text{minimize} \quad -\frac{4000}{(10+R)^2}$$
$$\text{subject to} \quad -R \leq 0$$

The derivative of the objective function is $\dfrac{8000}{(10+R)^3}$
Thus, the KKT condition is

$$\frac{8000}{(10+R)^3} - \mu = 0$$
$$\mu \geq 0, \quad \mu R = 0, \quad -R \leq 0$$

Example 21.2

- As before, we consider two cases. In the first case, suppose that $\mu > 0$. Then, $R = 0$, which is feasible. For the second case, suppose that $\mu = 0$. But this contradicts the first condition. Therefore, the only solution to the KKT condition is $R = 0$, $\mu = 0$

Karush-Kuhn-Tucker Condition

- In the case when the objective function is to be maximized, that is, when the optimization problem has the form

$$\begin{aligned} \text{maximize } & f(\boldsymbol{x}) \\ \text{subject to } & \boldsymbol{h}(\boldsymbol{x}) = \boldsymbol{0} \\ & \boldsymbol{g}(\boldsymbol{x}) \leq \boldsymbol{0} \end{aligned}$$

The KKT condition can be written as

- 1. $\boldsymbol{\mu}^* \geq \boldsymbol{0}$
- 2. $-Df(\boldsymbol{x}^*) + \boldsymbol{\lambda}^{*T} D\boldsymbol{h}(\boldsymbol{x}^*) + \boldsymbol{\mu}^{*T} D\boldsymbol{g}(\boldsymbol{x}^*) = \boldsymbol{0}^T$
- 3. $\boldsymbol{\mu}^{*T}\boldsymbol{g}(\boldsymbol{x}^*) = 0$
- 4. $\boldsymbol{h}(\boldsymbol{x}^*) = \boldsymbol{0}$
- 5. $\boldsymbol{g}(\boldsymbol{x}^*) \leq \boldsymbol{0}$

Karush-Kuhn-Tucker Condition

- The above is easily derived by converting the maximization problem above into a minimization problem, by multiplying the objective function by -1. It can further rewritten as
  - 1. $\boldsymbol{\mu}^* \leq \mathbf{0}$
  - 2. $Df(\boldsymbol{x}^*) + \boldsymbol{\lambda}^{*T} D\boldsymbol{h}(\boldsymbol{x}^*) + \boldsymbol{\mu}^{*T} D\boldsymbol{g}(\boldsymbol{x}^*) = \mathbf{0}^T$
  - 3. $\boldsymbol{\mu}^{*T}\boldsymbol{g}(\boldsymbol{x}^*) = 0$
  - 4. $\boldsymbol{h}(\boldsymbol{x}^*) = \mathbf{0}$
  - 5. $\boldsymbol{g}(\boldsymbol{x}^*) \leq \mathbf{0}$
- The form shown above is obtained from the preceding one by changing the signs of $\boldsymbol{\mu}^*$ and $\boldsymbol{\lambda}^*$ and multiplying condition 2 by -1.

Karush-Kuhn-Tucker Condition

- We can simply derive the KKT condition for the case when the inequality constraint is of the form $\boldsymbol{g}(\boldsymbol{x}) \geq \mathbf{0}$. Specifically, consider the problem

$$\begin{aligned} &\text{minimize } f(\boldsymbol{x}) \\ &\text{subject to } \boldsymbol{h}(\boldsymbol{x}) = \mathbf{0} \\ &\qquad\qquad \boldsymbol{g}(\boldsymbol{x}) \geq \mathbf{0} \end{aligned}$$

- We multiply the inequality constraint function by -1 to obtain $-\boldsymbol{g}(\boldsymbol{x}) \leq \mathbf{0}$. Thus, the KKT condition for this case is
  - 1. $\boldsymbol{\mu}^* \geq \mathbf{0}$
  - 2. $Df(\boldsymbol{x}^*) + \boldsymbol{\lambda}^{*T} D\boldsymbol{h}(\boldsymbol{x}^*) - \boldsymbol{\mu}^{*T} D\boldsymbol{g}(\boldsymbol{x}^*) = \mathbf{0}^T$
  - 3. $\boldsymbol{\mu}^{*T}\boldsymbol{g}(\boldsymbol{x}^*) = 0$
  - 4. $\boldsymbol{h}(\boldsymbol{x}^*) = \mathbf{0}$
  - 5. $\boldsymbol{g}(\boldsymbol{x}^*) \geq \mathbf{0}$

Karush-Kuhn-Tucker Condition

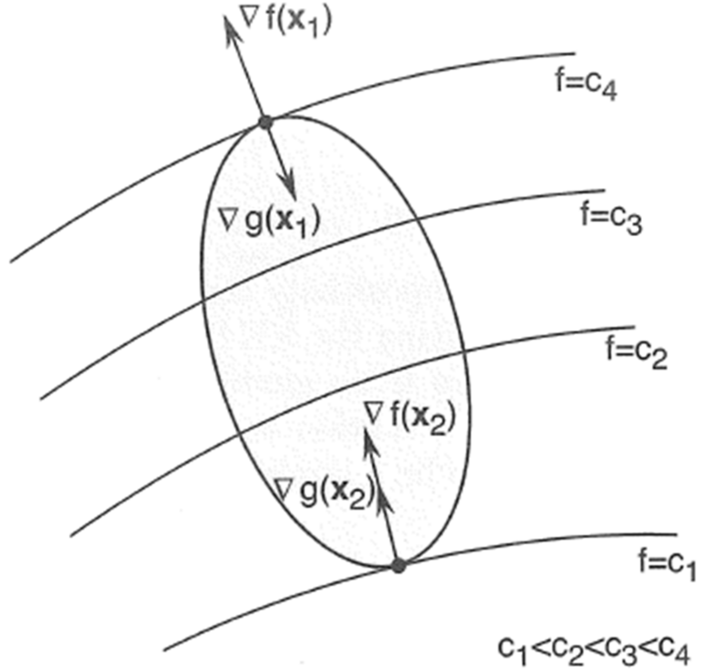
- Changing the sign of $\boldsymbol{\mu}^*$ as before, we obtain
  - 1. $\boldsymbol{\mu}^* \leq \mathbf{0}$
  - 2. $Df(\boldsymbol{x}^*) + \boldsymbol{\lambda}^{*T} D\boldsymbol{h}(\boldsymbol{x}^*) + \boldsymbol{\mu}^{*T} D\boldsymbol{g}(\boldsymbol{x}^*) = \mathbf{0}^T$
  - 3. $\boldsymbol{\mu}^{*T}\boldsymbol{g}(\boldsymbol{x}^*) = 0$
  - 4. $\boldsymbol{h}(\boldsymbol{x}^*) = \mathbf{0}$
  - 5. $\boldsymbol{g}(\boldsymbol{x}^*) \geq \mathbf{0}$
- For the problem

$$\begin{aligned} \text{maximize } & f(\boldsymbol{x}) \\ \text{subject to } & \boldsymbol{h}(\boldsymbol{x}) = \mathbf{0} \\ & \boldsymbol{g}(\boldsymbol{x}) \geq \mathbf{0} \end{aligned}$$

  the KKT condition is exactly the same as in Theorem 21.1, except for the reversal of the inequality constraint.

Example 21.3

- In Figure 21.3, the two points $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$ are feasible points; that is, $g(\boldsymbol{x}_1) \geq 0$ and $g(\boldsymbol{x}_2) \geq 0$, and they satisfy the KKT condition. The point $\boldsymbol{x}_1$ is a maximizer. The KKT condition for this point (with KKT multiplier μ_1) is
  - 1. $\mu_1 \geq 0$
  - 2. $\nabla f(\boldsymbol{x}_1) + \mu_1 \nabla g(\boldsymbol{x}_1) = \mathbf{0}$
  - 3. $\mu_1 g(\boldsymbol{x}_1) = 0$
  - 4. $g(\boldsymbol{x}_1) \geq 0$
- Figure 21.3



Example 21.3

- The point $\boldsymbol{x}_2$ is a minimizer of f. The KKT condition for this point (with KKT multiplier μ_2) is
  - 1. $\mu_2 \leq 0$
  - 2. $\nabla f(\boldsymbol{x}_2) + \mu_2 \nabla g(\boldsymbol{x}_2) = \mathbf{0}$
  - 3. $\mu_2 g(\boldsymbol{x}_2) = 0$
  - 4. $g(\boldsymbol{x}_2) \geq 0$

Example 21.4

- Consider the problem

$$\begin{aligned} &\text{minimize } f(x_1, x_2) \\ &\text{subject to } x_1, x_2 \geq 0 \end{aligned}$$

  where $f(x_1, x_2) = x_1^2 + x_2^2 + x_1 x_2 - 3x_1$

  The KKT condition for this problem is

  - 1. $\boldsymbol{\mu} = [\mu_1, \mu_2]^T \leq \mathbf{0}$
  - 2. $Df(\boldsymbol{x}) + \boldsymbol{u}^T = \mathbf{0}^T$
  - 3. $\boldsymbol{\mu}^T \boldsymbol{x} = 0$
  - 4. $\boldsymbol{x} \geq 0$

- We have $Df(\boldsymbol{x}) = [2x_1 + x_2 - 3, x_1 + 2x_2]$. This gives

$$\begin{aligned} 2x_1 + x_2 + \mu_1 &= 3 \\ x_1 + 2x_2 + \mu_2 &= 0 \\ \mu_1 x_1 + \mu_2 x_2 &= 0 \end{aligned}$$

Example 21.4

- We now have four variables, three equations, and the inequality constraints on each variable. To find a solution $(\boldsymbol{x}^*, \boldsymbol{\mu}^*)$, we first try $\mu_1^* = 0, x_2^* = 0$, which gives $x_1^* = \frac{3}{2}, \mu_2^* = -\frac{3}{2}$ The above satisfies all the KKT and feasibility conditions.
- In a similar fashion, we can try $\mu_2^* = 0, x_1^* = 0$, which gives $x_2^* = 0, \mu_1^* = 3$. This point clearly violates the nonpositivity constraints on μ_1^* .
- The feasible point above satisfying the KKT condition is only a candidate for a minimizer. However, there is no guarantee that the point is indeed a minimizer, because the KKT condition is, in general, only necessary. A sufficient condition for a point to be a minimizer is given as follows.

Karush-Kuhn-Tucker Condition

- Example 21.4 is a special case of a more general problem of the form

$$\text{minimize } f(\boldsymbol{x})$$
$$\text{subject to } \boldsymbol{x} \geq \mathbf{0}$$

  The KKT condition for this problem has the form

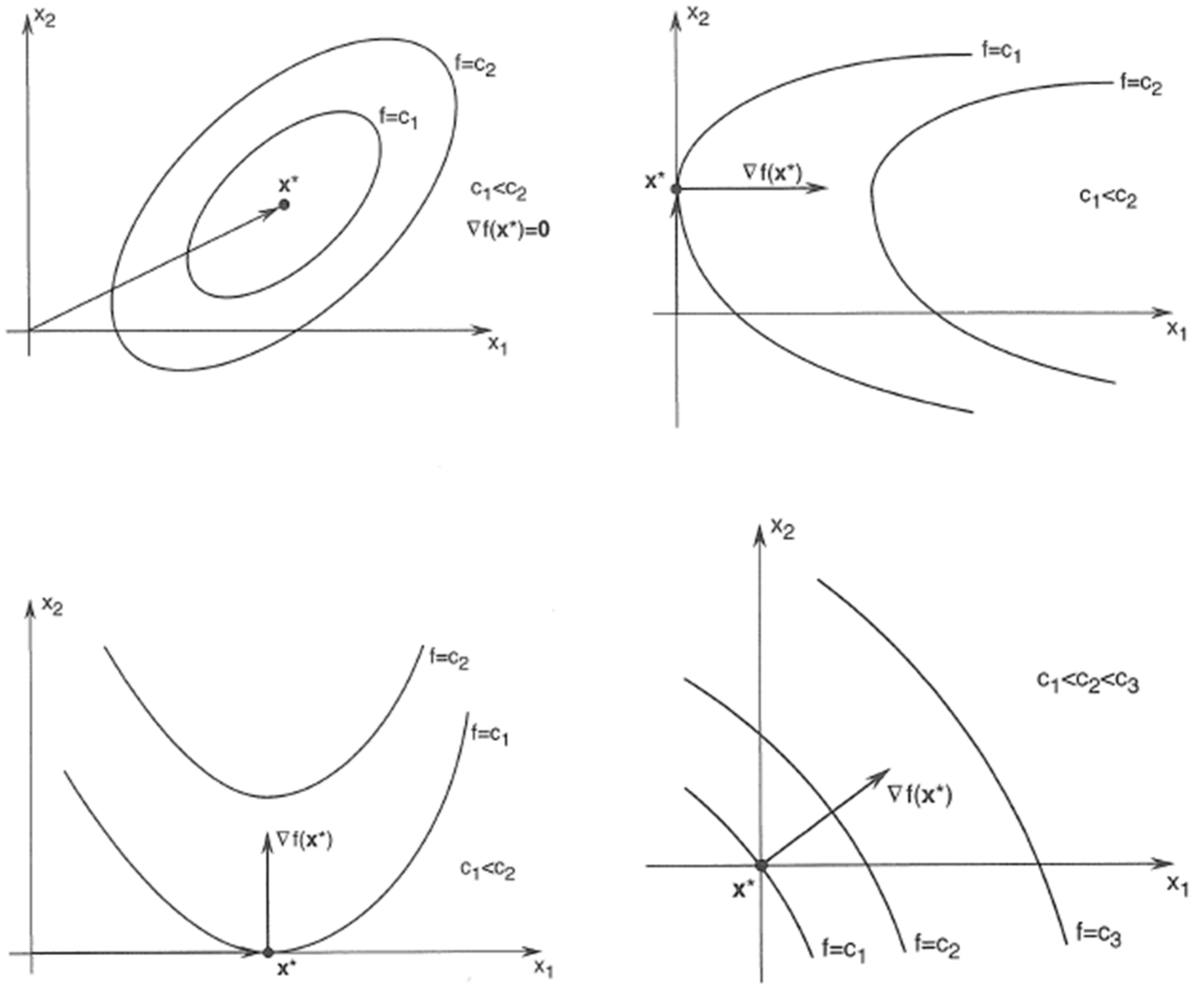
$$\boldsymbol{\mu} \leq \mathbf{0} \qquad \boldsymbol{\mu}^T \boldsymbol{x} = 0$$
$$\nabla f(\boldsymbol{x}) + \boldsymbol{\mu} = \mathbf{0} \qquad \boldsymbol{x} \geq \mathbf{0}$$

- For the above, we can eliminate $\boldsymbol{\mu}$ to obtain

$$\nabla f(\boldsymbol{x}) \geq \mathbf{0} \qquad \boldsymbol{x} \geq \mathbf{0}$$
$$\boldsymbol{x}^T \nabla f(\boldsymbol{x}) = 0$$

- Some possible points in $\mathbb{R}^2$ that satisfy these conditions are depicted in Figure 21.4.

Karush-Kuhn-Tucker Condition



Second-Order Conditions

- We can also give second-order necessary and sufficient conditions for extremum problems involving inequality constraints. Define the following matrix:

$$\boldsymbol{L}(\boldsymbol{x}, \boldsymbol{\lambda}, \boldsymbol{\mu}) = \boldsymbol{F}(\boldsymbol{x}) + [\boldsymbol{\lambda}\boldsymbol{H}(\boldsymbol{x})] + [\boldsymbol{\mu}\boldsymbol{G}(\boldsymbol{x})]$$

where $\boldsymbol{F}(\boldsymbol{x})$ is the Hessian matrix of f at $\boldsymbol{x}$, and the notation $[\boldsymbol{\lambda}\boldsymbol{H}(\boldsymbol{x})]$ represents

$$[\boldsymbol{\lambda}\boldsymbol{H}(\boldsymbol{x})] = \lambda_1 \boldsymbol{H}_1(\boldsymbol{x}) + \cdots + \lambda_m \boldsymbol{H}_m(\boldsymbol{x})$$

as before. Similarly, the notation $[\boldsymbol{\mu}\boldsymbol{G}(\boldsymbol{x})]$ represents

$$[\boldsymbol{\mu}\boldsymbol{G}(\boldsymbol{x})] = \mu_1 \boldsymbol{G}_1(\boldsymbol{x}) + \cdots + \mu_p \boldsymbol{G}_p(\boldsymbol{x})$$

where $\boldsymbol{G}_k(\boldsymbol{x})$ is the Hessian of g_k at $\boldsymbol{x}$, given by

$$\boldsymbol{G}_k(\boldsymbol{x}) = \begin{bmatrix} \frac{\partial^2 g_k}{\partial^2 x_1}(\boldsymbol{x}) & \cdots & \frac{\partial^2 g_k}{\partial x_n \partial x_1}(\boldsymbol{x}) \\ \vdots & & \vdots \\ \frac{\partial^2 g_k}{\partial x_1 \partial x_n}(\boldsymbol{x}) & \cdots & \frac{\partial^2 g_k}{\partial^2 x_n}(\boldsymbol{x}) \end{bmatrix}$$

Second-Order Conditions

- In the following theorem, we use
$$T(\boldsymbol{x}^*) = \{\boldsymbol{y} \in \mathbb{R}^n : D\boldsymbol{h}(\boldsymbol{x}^*)\boldsymbol{y} = \boldsymbol{0}, Dg_j(\boldsymbol{x}^*)\boldsymbol{y} = 0, j \in J(\boldsymbol{x}^*)\}$$
that is, the tangent space to the surface defined by active constraints.
- Theorem 21.2. Second-Order Necessary Conditions. Let $\boldsymbol{x}^*$ be a local minimizer of $f : \mathbb{R}^n \to \mathbb{R}$ subject to $\boldsymbol{h}(\boldsymbol{x}) = \boldsymbol{0}, \boldsymbol{g}(\boldsymbol{x}) \leq \boldsymbol{0}$ $\boldsymbol{h} : \mathbb{R}^n \to \mathbb{R}^m, m \leq n, \boldsymbol{g} : \mathbb{R}^n \to \mathbb{R}^p$ and $f, \boldsymbol{h}, \boldsymbol{g} \in \mathcal{C}^2$. Suppose that $\boldsymbol{x}^*$ is regular. Then, there exist $\boldsymbol{\lambda}^* \in \mathbb{R}^m$ and $\boldsymbol{\mu}^* \in \mathbb{R}^p$ such that
  - 1. $\boldsymbol{\mu}^* \geq \boldsymbol{0}, Df(\boldsymbol{x}^*) + \boldsymbol{\lambda}^{*T} D\boldsymbol{h}(\boldsymbol{x}^*) + \boldsymbol{\mu}^{*T} D\boldsymbol{g}(\boldsymbol{x}^*) = \boldsymbol{0}^T, \boldsymbol{\mu}^{*T}\boldsymbol{g}(\boldsymbol{x}^*) = 0$
  - 2. For all $\boldsymbol{y} \in T(\boldsymbol{x}^*)$ we have $\boldsymbol{y}^T \boldsymbol{L}(\boldsymbol{x}^*, \boldsymbol{\lambda}^*, \boldsymbol{\mu}^*)\boldsymbol{y} \geq 0$

Second-Order Conditions

- We now state the second-order sufficient conditions for extremum problems involving inequality constraints. In the formulation of the result, we use the following set

$$\tilde{T}(\boldsymbol{x}^*, \boldsymbol{\mu}^*) = \{\boldsymbol{y} : D\boldsymbol{h}(\boldsymbol{x}^*)\boldsymbol{y} = \boldsymbol{0}, Dg_j(\boldsymbol{x}^*)\boldsymbol{y} = 0, j \in \tilde{J}(\boldsymbol{x}^*, \boldsymbol{\mu}^*)\}$$

where $\tilde{J}(\boldsymbol{x}^*, \boldsymbol{\mu}^*) = \{i, g_i(\boldsymbol{x}^*) = 0, \mu_i^* > 0\}$. Note that $\tilde{J}(\boldsymbol{x}^*, \boldsymbol{\mu}^*)$ is a subset of $J(\boldsymbol{x}^*)$. This, in turn, implies that $T(\boldsymbol{x}^*)$ is a subset of $\tilde{T}(\boldsymbol{x}^*, \boldsymbol{\mu}^*)$

Second-Order Conditions

- Theorem 21.3. Second-Order Sufficient Conditions. Suppose that $f, \boldsymbol{h}, \boldsymbol{g} \in \mathcal{C}^2$ and there exist a feasible point $\boldsymbol{x}^* \in \mathbb{R}^n$ and vectors $\boldsymbol{\lambda}^* \in \mathbb{R}^m$ and $\boldsymbol{\mu}^* \in \mathbb{R}^p$ such that
  - 1. $\boldsymbol{\mu}^* \geq \mathbf{0}, Df(\boldsymbol{x}^*) + \boldsymbol{\lambda}^{*T} D\boldsymbol{h}(\boldsymbol{x}^*) + \boldsymbol{\mu}^{*T} D\boldsymbol{g}(\boldsymbol{x}^*) = \mathbf{0}^T, \boldsymbol{\mu}^{*T}\boldsymbol{g}(\boldsymbol{x}^*) = 0$
  - 2. For all $\boldsymbol{y} \in \tilde{T}(\boldsymbol{x}^*, \boldsymbol{\mu}^*), \boldsymbol{y} \neq \mathbf{0}$, we have $\boldsymbol{y}^T \boldsymbol{L}(\boldsymbol{x}^*, \boldsymbol{\lambda}^*, \boldsymbol{\mu}^*)\boldsymbol{y} > 0$

  Then, $\boldsymbol{x}^*$ is a strict local minimizer of f subject to $\boldsymbol{h}(\boldsymbol{x}) = \mathbf{0}, \boldsymbol{g}(\boldsymbol{x}) \leq \mathbf{0}$

Example 21.5

- Consider the following problem:

$$\text{minimize } x_1 x_2$$
$$\text{subject to } x_1 + x_2 \geq 2 \qquad x_2 \geq x_1$$

  - a. Write down the KKT condition for this problem
  - Write $f(\boldsymbol{x}) = x_1 x_2, g_1(\boldsymbol{x}) = 2 - x_1 - x_2$, and $g_2(\boldsymbol{x}) = x_1 - x_2$. The KKT condition is

$$\begin{aligned}
&x_2 - \mu_1 + \mu_2 = 0, \\
&x_1 - \mu_1 - \mu_2 = 0, \\
&\mu_1(2 - x_1 - x_2) + \mu_2(x_1 - x_2) = 0, \\
&\mu_1, \mu_2 \geq 0, \\
&2 - x_1 - x_2 \leq 0, \\
&x_1 - x_2 \leq 0
\end{aligned}$$

$$T(\boldsymbol{x}^*) = \{\boldsymbol{y} \in \mathbb{R}^n : D\boldsymbol{h}(\boldsymbol{x}^*)\boldsymbol{y} = \boldsymbol{0}, Dg_j(\boldsymbol{x}^*)\boldsymbol{y} = 0, j \in J(\boldsymbol{x}^*)\}$$

Example 21.5

- b. Find all points (and KKT multipliers) satisfying the KKT condition. In each case, determine if the point is regular.
- It is easy to check that $\mu_1 \neq 0, \mu_2 \not> 0$. This leaves us with only one solution to the KKT condition: $x_1^* = x_2^* = 1, \mu_1^* = 1, \mu_2^* = 0$ For this point we have $Dg_1(\boldsymbol{x}^*) = [-1, -1]$ and $Dg_2(\boldsymbol{x}^*) = [1, -1]$ Hence, $\boldsymbol{x}^*$ is regular.

- c. Find all points in part b that also satisfy the SONC.
- Both constraints are active. Hence, because $\boldsymbol{x}^*$ is regular, $T(\boldsymbol{x}^*) = \{\boldsymbol{0}\}$. This implies that the SONC is satisfied.

Example 21.5

- d. Find all points in part c that also satisfy the SOSC.
- Now
$$\boldsymbol{L}(\boldsymbol{x}^*, \boldsymbol{\mu}^*) = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

Moreover, $\tilde{T}(\boldsymbol{x}^*, \boldsymbol{\mu}^*) = \{\boldsymbol{y} : [-1, -1]\boldsymbol{y} = 0\} = \{\boldsymbol{y} : y_1 = -y_2\}$. Pick $\boldsymbol{y} = [1, -1]^T \in \tilde{T}(\boldsymbol{x}^*, \boldsymbol{\mu}^*)$. We have $\boldsymbol{y}^T \boldsymbol{L}(\boldsymbol{x}^*, \boldsymbol{\mu}^*)\boldsymbol{y} = -2 < 0$, which means that the SOSC fails.

- e. Find all points in part c that are local minimizers.
- In fact, the point $\boldsymbol{x}^*$ is not a local minimizer. To see this, draw a picture of the constraint set and level sets of the objective function. Moving in the feasible direction $[1, 1]^T$, the objective function increases; but moving in the feasible direction $[-1, 1]^T$ the objective function decreases.

Example 21.6

- We wish to minimize $f(\boldsymbol{x}) = (x_1 - 1)^2 + x_2 - 2$ subject to

$$h(\boldsymbol{x}) = x_2 - x_1 - 1 = 0$$
$$g(\boldsymbol{x}) = x_1 + x_2 - 2 \leq 0$$

For all $\boldsymbol{x} \in \mathbb{R}^2$, we have $Dh(\boldsymbol{x}) = [-1, 1], \quad Dg(\boldsymbol{x}) = [1, 1]$
Thus, $\triangledown h(\boldsymbol{x})$ and $\triangledown g(\boldsymbol{x})$ are linearly independent and hence all feasible points are regular. We first write the KKT condition. Because $Df(\boldsymbol{x}) = [2x_1 - 2, 1]$

$$Df(\boldsymbol{x}) + \lambda Dh(\boldsymbol{x}) + \mu Dg(\boldsymbol{x}) = [2x_1 - 2 - \lambda + \mu, 1 + \lambda + \mu] = \mathbf{0}^T$$

$$\begin{aligned} &\mu(x_1 + x_2 - 2) = 0 \\ &\mu \geq 0, \\ &x_2 - x_1 - 1 = 0, \\ &x_1 + x_2 - 2 \leq 0 \end{aligned}$$

Example 21.6

- To find points that satisfy the conditions above, we first try $\mu > 0$, which implies that $x_1 + x_2 - 2 = 0$. Thus, we are faced with a system of four linear equations

$$\begin{aligned} &2x_1 - 2 - \lambda + \mu = 0, \\ &1 + \lambda + \mu = 0, \\ &x_2 - x_1 - 1 = 0, \\ &x_1 + x_2 - 2 = 0 \end{aligned}$$

Solving the system of equations above, we obtain

$$x_1 = \frac{1}{2}, x_2 = \frac{3}{2}, \lambda = -1, \mu = 0$$

However, the above is not a legitimate solution to the KKT condition, because we obtained $\mu = 0$, which contradicts the assumption that $\mu > 0$

Example 21.6

- In the second try, we assume that $\mu = 0$. Thus, we have to solve the system of equations

$$2x_1 - 2 - \lambda = 0,$$
$$1 + \lambda = 0,$$
$$x_2 - x_1 - 1 = 0$$

  and the solutions must satisfy $g(x_1, x_2) = x_1 + x_2 - 2 \leq 0$

- Solving the equations above, we obtain

$$x_1 = \frac{1}{2}, x_2 = \frac{3}{2}, \lambda = -1$$

  Note that $\boldsymbol{x}^* = [1/2, 3/2]^T$ satisfies the constraint $g(\boldsymbol{x}^*) \leq 0$. The point $\boldsymbol{x}^*$ satisfying the KKT necessary condition is therefore the candidate for being a minimizer.

Example 21.6

- We now verify if $\boldsymbol{x}^* = [1/2, 3/2]^T, \lambda^* = -1, \mu^* = 0$, satisfy the second-order sufficient conditions. For this, we form the matrix

$$\boldsymbol{L}(\boldsymbol{x}^*, \lambda^*, \mu^*) = \boldsymbol{F}(\boldsymbol{x}^*) + \lambda^* \boldsymbol{H}(\boldsymbol{x}^*) + \mu^* \boldsymbol{G}(\boldsymbol{x}^*)$$
$$= \begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix} + (-1) \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} + (0) \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$
$$= \begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix}$$

We then find the subspace $\tilde{T}(\boldsymbol{x}^*, \mu^*) = \{\boldsymbol{y} : Dh(\boldsymbol{x}^*)\boldsymbol{y} = 0\}$ Note that because $\mu^* = 0$, the active constraint $g(\boldsymbol{x}^*) = 0$ does not enter the computation of $\tilde{T}(\boldsymbol{x}^*, \mu^*)$. Note also that in this case, $T(\boldsymbol{x}^*) = \{\boldsymbol{0}\}$. We have

$$\tilde{T}(\boldsymbol{x}^*, \mu^*) = \{\boldsymbol{y} : [-1, 1]\boldsymbol{y} = 0\} = \{[a, a]^T : a \in \mathbb{R}\}$$

Example 21.6

- We then check for positive definiteness of $\boldsymbol{L}(\boldsymbol{x}^*, \lambda^*, \mu^*)$ on $\tilde{T}(\boldsymbol{x}^*, \mu^*)$. We have

$$\boldsymbol{y}^T \boldsymbol{L}(\boldsymbol{x}^*, \lambda^*, \mu^*)\boldsymbol{y} = [a, a] \begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} a \\ a \end{bmatrix} = 2a^2$$

  Thus, $\boldsymbol{L}(\boldsymbol{x}^*, \lambda^*, \mu^*)$ is positive definite on $\tilde{T}(\boldsymbol{x}^*, \mu^*)$. Observe that $\boldsymbol{L}(\boldsymbol{x}^*, \lambda^*, \mu^*)$ is, in fact, only positive semidefinite on $\mathbb{R}^2$

- By the second-order sufficient conditions, we conclude that $\boldsymbol{x}^* = [1/2, 3/2]^T$ is a strict local minimizer.