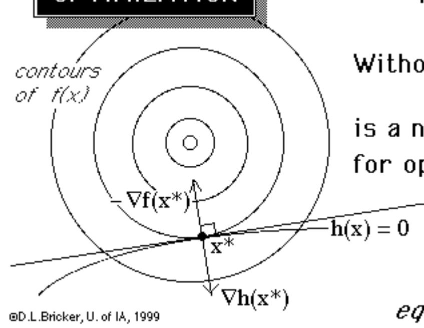
Optimality Conditions

author

EQUALITY-CONSTRAINED OPTIMIZATION

Minimize $f(x)$
subject to $h(x) = 0$

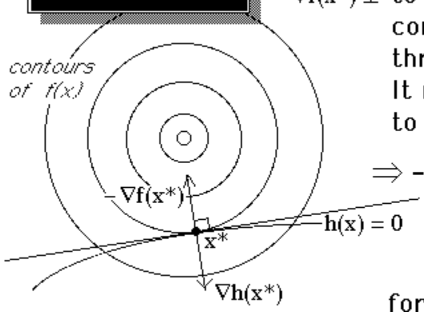
where $x=(x_1,x_2, \cdots x_n)$
f & h differentiable



Without the constraint,
$$\nabla f(x^*) = 0$$
is a necessary condition for optimality of x^*.

What is the corresponding condition with equality constraint?

EQUALITY-CONSTRAINED OPTIMIZATION



$-\nabla f(x^*)$ is the steepest descent direction

$-\nabla f(x^*) \perp$ to the tangent of the contour of f passing through x^*

It must also be $\perp$ to the curve $h(x)=0$!

$\Rightarrow -\nabla f(x^*)$ & $\nabla h(x^*)$ must be collinear

$$-\nabla f(x^*) = \lambda \, \nabla h(x^*)$$

for some value of λ

LAGRANGIAN FUNCTION

Define the Lagrangian function
$$L(x,\lambda) \equiv f(x) + \lambda\, h(x)$$

Then the necessary condition for optimality may be written conveniently as

i.e., $\nabla_{x,\lambda} L(x^*,\lambda^*) = 0$

$$\begin{cases} \dfrac{\partial L(x^*,\lambda^*)}{\partial x_i} = \dfrac{\partial f(x^*)}{\partial x_i} + \lambda \dfrac{\partial h(x^*)}{\partial x_i} = 0 \\ \dfrac{\partial L(x^*,\lambda^*)}{\partial \lambda} = h(x^*) = 0 \end{cases} \Rightarrow \begin{cases} -\nabla f(x^*) = \lambda^* \nabla h(x^*) \\ h(x^*) = 0 \end{cases}$$

EXAMPLE

Minimize $\frac{1}{2}x_1^2 + \frac{1}{2}x_2^2 - x_1x_2 - 3x_2$
s.t. $x_1 + x_2 = 3$,
i.e., $h(x) = x_1 + x_2 - 3 = 0$

Lagrangian function is
$$L(x_1,x_2,\lambda) = \tfrac{1}{2}x_1^2 + \tfrac{1}{2}x_2^2 - x_1x_2 - 3x_2 + \lambda(x_1 + x_2 - 3)$$

Necessary conditions for optimality of x^*,λ^*

$$\begin{cases} \dfrac{\partial L(x^*,\lambda^*)}{\partial x_1} = x_1^* - x_2^* + \lambda^* = 0 \\ \dfrac{\partial L(x^*,\lambda^*)}{\partial x_2} = x_2^* - x_1^* - 3 + \lambda^* = 0 \\ \dfrac{\partial L(x^*,\lambda^*)}{\partial \lambda} = x_1^* + x_2^* - 3 = 0 \end{cases}$$

These conditions have a unique solution:

$$\begin{cases} x_1^* - x_2^* + \lambda^* = 0 \\ x_2^* - x_1^* - 3 + \lambda^* = 0 \\ x_1^* + x_2^* - 3 = 0 \end{cases} \Rightarrow \begin{cases} x_1^* = 3/4 \\ x_2^* = 9/4 \\ \lambda^* = 3/2 \end{cases}$$

K-K-T Point

That is, if an optimal solution exists, it must be

$$\begin{cases} x_1^* = 3/4 \\ x_2^* = 9/4 \end{cases}$$

A K-K-T point is not necessarily optimal!

INEQUALITY-CONSTRAINED OPTIMIZATION

Minimize $f(x)$
subject to $g(x) \le 0$

where $x=(x_1,x_2, \cdots x_n)$
f & g differentiable

Can we convert to an equality-constrained optimization problem?

Suppose we add a "slack" variable, as in LP:

$$\begin{cases} g(x) + s = 0 \\ s \ge 0 \end{cases}$$

We must also add a nonnegativity constraint, thus trading one inequality for another!

INEQUALITY-CONSTRAINED OPTIMIZATION

Minimize $f(x)$
subject to $g(x) \le 0$

where $x=(x_1,x_2, \cdots x_n)$
f & g differentiable

Instead, let's use $\frac{1}{2}s^2$ for the slack...
...obviously, we needn't include $s^2 \ge 0$ as an additional constraint!

$$g(x) + \tfrac{1}{2}s^2 = 0$$

What are now the necessary conditions for optimality (with $h(x) = g(x) + \frac{1}{2}s^2 = 0$)?

$$\nabla_{x,s,\lambda} L(x^*,s^*,\lambda^*) = 0$$

where

$$L(x,s,\lambda) \equiv f(x) + \lambda\left[g(x)+\tfrac{1}{2}s^2\right]$$

i.e.,

$$\begin{cases} \dfrac{\partial L(x^*,s^*,\lambda^*)}{\partial x_i} = \dfrac{\partial f(x^*)}{\partial x_i} + \lambda\dfrac{\partial g(x^*)}{\partial x_i} = 0 \\ \dfrac{\partial L(x^*,s^*,\lambda^*)}{\partial \lambda} = g(x^*) + \tfrac{1}{2}(s^*)^2 = 0 \\ \dfrac{\partial L(x^*,s^*,\lambda^*)}{\partial s} = \lambda^* s^* = 0 \end{cases}$$

$$\begin{cases} \dfrac{\partial L(x^*,s^*,\lambda^*)}{\partial x_i} = \dfrac{\partial f(x^*)}{\partial x_i} + \lambda\dfrac{\partial g(x^*)}{\partial x_i} = 0 \\ \dfrac{\partial L(x^*,s^*,\lambda^*)}{\partial \lambda} = g(x^*) + \tfrac{1}{2}(s^*)^2 = 0 \\ \dfrac{\partial L(x^*,s^*,\lambda^*)}{\partial s} = \lambda^* s^* = 0 \end{cases} \Leftrightarrow \begin{cases} -\nabla f(x^*) = \lambda \nabla g(x^*) \\ g(x^*) \le 0 \\ \lambda^* g(x^*) = 0 \end{cases}$$

"complementary slackness"

Either the inequality is tight or the Lagrange multiplier is zero (or both)

EXAMPLE

Minimize $x_1^2 + \frac{1}{2}x_2^2 - 8x_1 - 2x_2 - 60$

subject to $40x_1 + 20x_2 \le 140$

$$L(x,s,\lambda) = x_1^2 + \tfrac{1}{2}x_2^2 - 8x_1 - 2x_2 - 60 + \lambda\left[40x_1 + 20x_2 + \tfrac{1}{2}s^2 - 140\right]$$

$$\begin{cases} \dfrac{\partial L(x,\lambda)}{\partial x_1} = 2x_1 - 8 + 40\lambda = 0 \\ \dfrac{\partial L(x,\lambda)}{\partial x_2} = x_2 - 2 + 20\lambda = 0 \\ \dfrac{\partial L(x,\lambda)}{\partial \lambda} = 40x_1 + 20x_2 + \tfrac{1}{2}s^2 - 140 = 0 \\ \dfrac{\partial L(x,\lambda)}{\partial s} = \lambda s = 0 \end{cases}$$

Optimality conditions

Note that this is a nonlinear system!

Rather than to attempt a direct solution, using the Newton-Raphson method, for example, we consider TWO cases, based upon the complementary slackness constraint:

$$\lambda s = 0$$

Case 1: $\lambda = 0$

Case 2: $s = 0$

We then seek a solution for each case.

Case 1: $\lambda = 0$

$2x_1 - 8 = 0 \Rightarrow x_1 = 4$

$x_2 - 2 = 0 \Rightarrow x_2 = 2$

$\Rightarrow 40x_1 + 20x_2 + \frac{1}{2}s^2 - 140 = 0 \Rightarrow s^2 = -120$

Clearly, this system of equations has no feasible solution!

Case 2: $s = 0$

$$\begin{cases} 2x_1 + 40\lambda = 8 \\ x_2 + 20\lambda = 2 \\ 40x_1 + 20x_2 = 140 \end{cases}$$

This is a linear system, which may be solved by Gauss elimination:

$x_1 = 3$
$x_2 = 1$
$\lambda = 0.05$

K-K-T Point

That is, this is the *only* solution of the necessary conditions for optimality, so *if* the original problem has a solution, this must be it!

Generalizing to Several Constraints....

EQUALITIES

Minimize $f(x_1,x_2,\cdots x_n)$

subject to $\begin{cases} h_1(x_1,x_2,\cdots x_n) = 0 \\ h_2(x_1,x_2,\cdots x_n) = 0 \\ \vdots \\ h_m(x_1,x_2,\cdots x_n) = 0 \end{cases}$

INEQUALITIES

Minimize $f(x_1,x_2,\cdots x_n)$

subject to $\begin{cases} g_1(x_1,x_2,\cdots x_n) \le 0 \\ g_2(x_1,x_2,\cdots x_n) \le 0 \\ \vdots \\ g_m(x_1,x_2,\cdots x_n) \le 0 \end{cases}$

EQUALITIES

$$\begin{cases} \text{Minimize } f(x_1,x_2,\cdots x_n) \\ \text{subject to } h_i(x^*) = 0 \quad \forall\, i=1,\ldots m \end{cases}$$

Define the Lagrangian function

$$L(x_1,\cdots x_n,\lambda_1,\cdots \lambda_m) = f(x) + \sum_{i=1}^{m} \lambda_i h_i(x)$$

with one multiplier per equality constraint.

$$\nabla_{x,\lambda} L(x^*,\lambda^*) = 0 \quad \Rightarrow \quad \begin{cases} -\nabla f(x^*) = \sum_{i=1}^{m} \lambda_i^* \nabla h_i(x^*) \\ h_i(x^*) = 0 \quad \forall\, i=1,\ldots m \end{cases}$$

INEQUALITIES

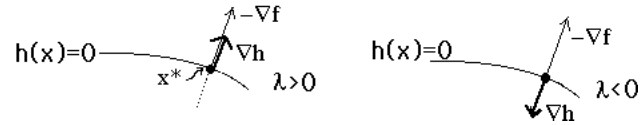
$$\begin{cases} \text{Minimize } f(x_1, x_2, \cdots x_n) \\ \text{subject to } g_i(x) \le 0 \quad \forall\, i=1,\ldots m \end{cases}$$

Introduce squared slack variables and define

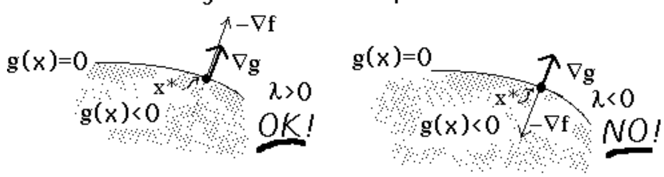
$$L(x_1, \cdots x_n, s_1, \cdots s_m, \lambda_1, \cdots \lambda_m) = f(x) + \sum_{i=1}^{m} \lambda_i \left[g_i(x) + \tfrac{1}{2} s_i^2 \right]$$

$$\nabla_{x,s,\lambda} L(x,s,\lambda) = 0 \Rightarrow \begin{cases} -\nabla f(x^*) = \sum_{i=1}^{m} \lambda_i^* \nabla g_i(x^*) \\ g_i(x_1, x_2, \cdots x_n) + \tfrac{1}{2} s_i^2 = 0 \quad \forall\, i=1,\ldots m \\ \lambda_i s_i = 0 \quad \forall\, i=1,\ldots m \end{cases}$$

In the case of equality constraints, the corresponding multiplier may be either positive or negative.



In the case of inequality constraints, we can deduce the sign of the multiplier!



The gradient of g will point outside of the feasible region of the inequality constraint $g(x) \le 0$, which must also be true of the steepest descent direction, if x^* is optimal!

Farkas' Lemma applied to NLP

Consider the problem

Minimize $f(x)$
subject to $g_i(x) \le 0,\ i=1,2,\ldots m$

Denote

$b \equiv -\nabla f(x^*)$

$A^i \equiv \nabla g_i(x^*)$

$y \equiv d$ *(direction vector)*

$x_i \equiv \lambda_i$ for $i \in I \equiv \{ i \mid g_i(x^*) = 0 \}$

(Lagrange multiplier) *(index set of tight constraints)*

Farkas' Lemma

1 $y^T A \le 0 \Rightarrow y^T b \le 0$

&

2 $\exists\, x$ such that $A x = b,\ x \ge 0$

are equivalent statements

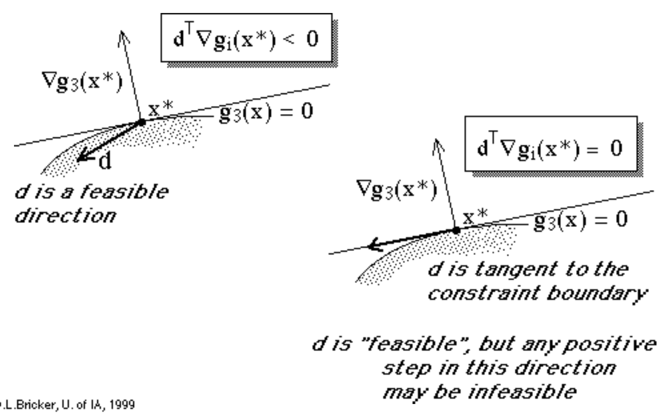
That is,

1 $d^T \nabla g_i(x^*) \le 0 \ \forall i \in I \Rightarrow -d^T \nabla f(x^*) \le 0$

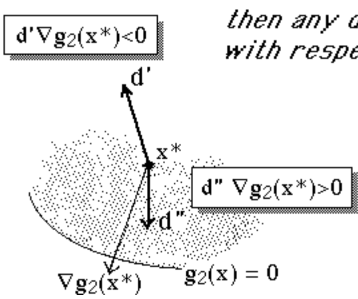
&

2 $\exists\, \lambda_i \ge 0$ such that $\sum_{i \in I} \lambda_i \nabla g_i(x^*) = -\nabla f(x^*)$

are equivalent statements



If a constraint is not tight, then any direction is feasible with respect to that constraint!



1 $d^T \nabla g_i(x^*) \le 0 \ \forall i \in I \Rightarrow -d^T \nabla f(x^*) \le 0$

directions satisfying $d^T \nabla g_i(x^*) \le 0 \ \forall i \in I$ are feasible directions

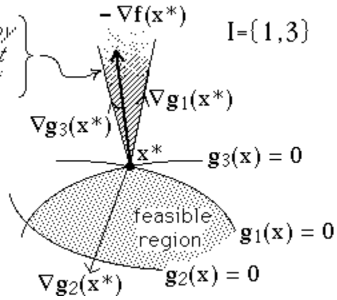
directions satisfying $d^T \nabla f(x^*) \ge 0$ are directions of ascent

And so statement #1 may be restated as:

1 Every feasible direction is non-improving

2 $\exists\ \lambda_i \geq 0$ such that $\sum_{i\in I} \lambda_i \nabla g_i(x^*) = -\nabla f(x^*)$

*cone generated by gradients of tight constraints at x**



Steepest descent direction is in the cone generated by gradients of tight constraints

Karush-Kuhn-Tucker Conditions

Necessary Condition for Optimality

If x^* is an optimal solution to

Minimize $f(x)$
subject to $g_i(x) \leq 0$, $i=1,2,\ldots m$

then

The directional derivative of $f(x)$ is nonnegative in every feasible direction at x^*

K-K-T "Necessary" Condition for Optimality

If x^* is an optimal solution to

Minimize $f(x)$
subject to $g_i(x) \leq 0$, $i=1,2,\ldots m$

then

The steepest descent direction at x^* is in the cone generated by the gradients of the tight constraints at x^*

Equivalent condition, according to Farkas' lemma

Is it true that if x^* is optimal, then every direction d which is "feasible", i.e., satisfying

$$d^T \nabla g_i(x^*) \leq 0\ \forall i \in I$$

cannot be a direction of descent, i.e.,

$$d^T \nabla f(x^*) \geq 0 \quad ?$$

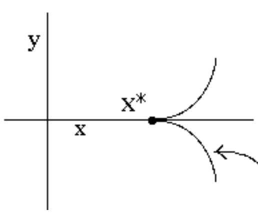
The next example will illustrate a case in which the KKT conditions are NOT satisfied by the optimal solution !

EXAMPLE

Find the point on the curve

$$y^2 - (x-1)^3 = 0$$

which is nearest the origin.

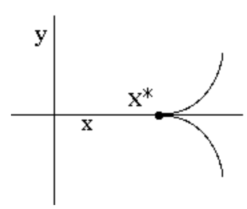


graph of g(x,y) = 0

Minimize $f(x,y) = \sqrt{x^2 + y^2}$
subject to
$g(x,y) = y^2 - (x-1)^3 = 0$

By inspection, we see that the optimal solution is $x^* = (1,0)$

Does it necessarily then satisfy the K-K-T conditions?



Define the Lagrangian function:

$$L(x,y,\lambda) = (x^2 + y^2)^{1/2} + \lambda\left[y^2 - (x-1)^3\right]$$

Stationary point of the Lagrangian function must satisfy:

$$\begin{cases} \dfrac{\partial L}{\partial x} = x\,(x^2+y^2)^{-1/2} - 3\lambda(x-1)^2 = 0 \\ \dfrac{\partial L}{\partial y} = y\,(x^2+y^2)^{-1/2} + 2\lambda y = 0 \\ \dfrac{\partial L}{\partial \lambda} = y^2 - (x-1)^3 = 0 \end{cases}$$

Does $x^*=(1,0)$ satisfy these conditions for some λ^* ?

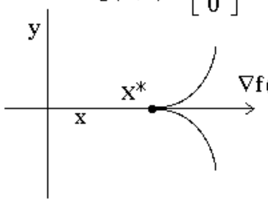
$$\begin{cases} \dfrac{\partial L}{\partial x} = 1 - \lambda \times 0 = 0 \\ \dfrac{\partial L}{\partial y} = 0 + \lambda \times 0 = 0 \\ \dfrac{\partial L}{\partial \lambda} = 0 + 0 = 0 \end{cases}$$

These equations cannot be satisfied by any $(x,y,\lambda) = (1,0,\lambda)$!

Does this mean that the K-K-T conditions are NOT, in fact, "necessary"?????

The cone generated by the gradient of the (tight) constraint is the null vector, which does not contain $\nabla f(x^*)$:

$$\nabla g(1,0) = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$



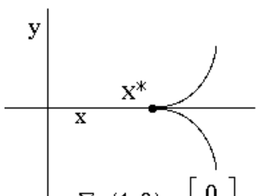
$$\nabla f(1,0) = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

$$\nabla f(1,0) \neq \lambda \nabla g(1,0) = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

for any λ

In a trivial sense, the set consisting of the single vector [0,0] is linearly dependent, since

$$k \begin{bmatrix} 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad \textit{for } k \neq 0!$$



$$\nabla g(1,0) = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

This example does not satisfy a constraint qualification!

Constraint Qualifications

Each of the following is sufficient to guarantee that

$$\{\, d \mid d^T \nabla g_i(x^*) \leq 0 \ \forall\, i \in I,\ d^T \nabla f(x^*) < 0 \,\} = \varnothing$$

for the local (or global) minimum x^*:

- All constraint functions are linear (Karlin)
- All constraint functions are convex and the feasible region has a nonempty interior (Karlin)
- The gradients of the binding constraints are linearly independent (Fiacco & McCormick)

Suppose that $\hat{x}$ is a locally optimal solution of the problem:

$$\begin{aligned} &\text{Minimize } f(x) \\ &\text{s.t. } g_i(x) \leq 0 \ \forall\, i = 1, \cdots m \\ &\qquad h_i(x) = 0 \ \forall\, i = 1, \cdots p \\ &\qquad x \in X \end{aligned}$$

Let $I = \{\, i \mid g_i(\hat{x}) = 0 \,\}$ be the indices of the tight constraints at $\hat{x}$

Finally, suppose that a "constraint qualification" is valid, e.g., $\{\nabla g_i(\hat{x}),\ i \in I\} \cup \{\nabla h_i(\hat{x}), 1 \leq i \leq p\}$ is a linearly independent set of vectors & g_i for $i \notin I$ is continuous at $\hat{x}$

Then there exists $\hat{\lambda}$ and $\hat{\mu}$ satisfying

$$\begin{aligned} &-\nabla f(\hat{x}) = \sum_{i=1}^{m} \hat{\lambda}_i \nabla g_i(\hat{x}) + \sum_{i=1}^{p} \hat{\mu}_i \nabla h_i(\hat{x}) \\ &\hat{\lambda}_i\, g_i(\hat{x}) = 0 \ \forall\, i = 1, \cdots m \\ &g_i(x) \leq 0 \ \forall\, i = 1, \cdots m \\ &h_i(x) = 0 \ \forall\, i = 1, \cdots p \\ &\hat{\lambda}_i \geq 0 \ \forall\, i = 1, \cdots m \\ &\hat{x} \in X \end{aligned}$$

K-K-T Conditions for Nonnegative Variables

$$\begin{aligned} &\text{Minimize } f(x) \\ &\text{s.t. } g_i(x) \leq 0 \ \forall\, i = 1, \cdots m \\ &\qquad x_j \geq 0,\ j = 1, \cdots p \end{aligned}$$

Writing all inequalities as "≤" :

$$\begin{aligned} &\text{Minimize } f(x) \\ &\text{s.t. } g_i(x) \leq 0 \ \forall\, i = 1, \cdots m \\ &\qquad -x_j \leq 0,\ j = 1, \cdots p \end{aligned}$$

Define the Lagrangian function:

$$L'(x, \lambda, \mu) = f(x) + \sum_{i=1}^{m} \lambda_i\, g_i(x) + \sum_{j=1}^{n} \mu_j\,(-x_j)$$

Equate the gradient of L' to zero:

$$\frac{\partial L'}{\partial x_j} = \frac{\partial f(x)}{\partial x_j} + \sum_{i=1}^{m} \lambda_i \frac{\partial g_i(x)}{\partial x_j} - \mu_j = 0 \quad \forall\, j \qquad \boxed{1}$$

$$\lambda_i\, g_i(x) = 0 \quad \forall\, i \qquad \boxed{2}$$

$$\mu_j\, x_j = 0 \quad \forall\, j \qquad \boxed{3}$$

$$\lambda_i \geq 0 \quad \forall\, i \qquad \boxed{4}$$

$$\mu_j \geq 0 \quad \forall\, j \qquad \boxed{5}$$

By solving equation **1** for μ_j and then substituting into conditions **3** and **5**, we can eliminate the Lagrange multipliers for the nonnegativity constraints:

$$\mu_j = \frac{\partial f(x)}{\partial x_j} + \sum_{i=1}^{m} \lambda_i \frac{\partial g_i(x)}{\partial x_j} \quad \forall\, j \qquad \boxed{1}$$

$$x_j \left[\frac{\partial f(x)}{\partial x_j} + \sum_{i=1}^{m} \lambda_i \frac{\partial g_i(x)}{\partial x_j} \right] = 0 \quad \forall\, j \qquad \boxed{3}$$

$$\frac{\partial f(x)}{\partial x_j} + \sum_{i=1}^{m} \lambda_i \frac{\partial g_i(x)}{\partial x_j} \geq 0 \quad \forall\, j \qquad \boxed{5}$$

If we form the Lagrangian as follows, omitting the nonnegativity constraints,

$$L(x,\lambda) = f(x) + \sum_{i=1}^{m} \lambda_i g_i(x)$$

then the K-K-T conditions may be stated as:

$$\frac{\partial L}{\partial x_j} = \frac{\partial f(x)}{\partial x_j} + \sum_{i=1}^{m} \lambda_i \frac{\partial g_i(x)}{\partial x_j} \quad \forall j \quad \textbf{1}$$

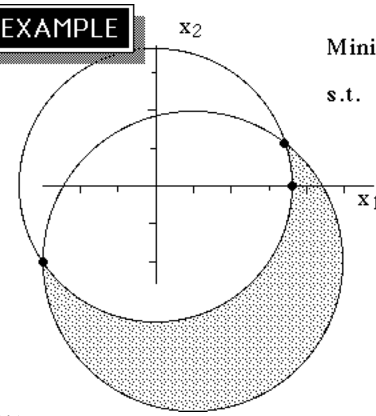
$$\lambda_i g_i(x) = 0 \quad \forall i \quad \textbf{2}$$

$$x_j \times \frac{\partial L}{\partial x_j} = 0 \quad \forall j \quad \textbf{3}$$

$$\lambda_i \geq 0 \quad \forall i \quad \textbf{4}$$

$$\frac{\partial L}{\partial x_j} \geq 0 \quad \forall j \quad \textbf{5}$$

EXAMPLE



Minimize x_1

s.t. $(x_1 - 1)^2 + (x_2 + 2)^2 \leq 16$

$x_1^2 + x_2^2 \geq 13$

i.e., x must lie inside circle which is centered at (1,-2) and outside circle centered at the origin.

Lagrangian Function:

$$L(x_1,x_2,\lambda_1,\lambda_2) = x_1 + \lambda_1\left[(x_1 - 1)^2 + (x_2 + 2)^2 - 16\right] + \lambda_2\left[13 - x_1^2 - x_2^2\right]$$

$$\frac{\partial L}{\partial x_1} = 1 + 2\lambda_1 (x_1 - 1) - 2\lambda_2 x_1 = 0$$

$$\frac{\partial L}{\partial x_1} = 2\lambda_1 (x_2 + 2) - 2\lambda_2 x_2 = 0$$

$$g_1(x) \leq 0 \ , \ g_2(x) \leq 0$$

$$\lambda_1 g_1(x) = 0 \ , \ \lambda_2 g_2(x) = 0$$

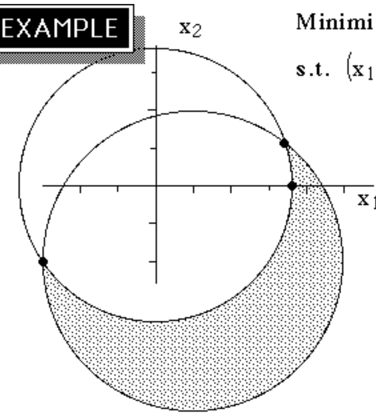
$$\lambda_1 \geq 0 \ , \ \lambda_2 \geq 0$$

Karush-Kuhn-Tucker Conditions

$$g_1(x) = (x_1-1)^2 + (x_2+2)^2 - 16$$

$$g_2(x) = 13 - x_1^2 - x_2^2$$

EXAMPLE



Minimize x_1

s.t. $(x_1 - 1)^2 + (x_2 + 2)^2 \leq 16$

$x_1^2 + x_2^2 \geq 13$

The points

(-3, -2) *global min.*

(3.4, 1.2) *local min.*

($\sqrt{13}$, 0)

satisfy the KKT conditions for optimality.

EXAMPLE

Minimize $-x_1 + x_2$

s.t.

$x_1^2 + x_2^2 - 2x_1 = 0$

$-1 \leq x_1 \leq 1$

$-1 \leq x_2 \leq 1$

Find the solution graphically, and test the K-K-T conditions.

Sufficiency of KKT conditions

Minimize $f(x)$

s.t. $g_i(x) \leq 0 \ \forall i = 1,\cdots m$

$h_i(x) = 0 \ \forall i = 1,\cdots p$

$x \in X$

Suppose that $(\hat{x},\hat{\lambda},\hat{\mu})$ satisfies the K-K-T conditions

$$-\nabla f(\hat{x}) = \sum_{i=1}^{m} \hat{\lambda}_i \nabla g_i(\hat{x}) + \sum_{i=1}^{p} \hat{\mu}_i \nabla h_i(\hat{x})$$

$$\hat{\lambda}_i g_i(\hat{x}) = 0 \ \forall i = 1,\cdots m$$

$$\hat{\lambda}_i \geq 0 \ \forall i = 1,\cdots m$$

$$\hat{x} \in X$$

K-K-T Sufficiency Conditions

Let $I = \{ i \mid g_i(\hat{x}) = 0 \}$ be the indices of the tight inequality constraints at $\hat{x}$

Suppose that
- f is convex
- g_i is convex for $i \in I$
- h_i is linear for $i = 1,\cdots p$

Then $\hat{x}$ is a globally optimal solution.

Quasiconvex Functions

A function g is quasiconvex if, for each x' and x",

$$g\left(\lambda x' + (1-\lambda)x''\right) \leq \text{maximum}\{g(x'), g(x'')\} \ \forall \lambda \in (0,1)$$

or, equivalently, if all of its *level sets* are convex:

$$L_C = \{ x \mid g(x) \leq C \} \text{ is convex } \forall C$$

Functions in inequality constraints $g_i(x) \leq 0$ need only be quasiconvex to get convex feasible region!

Pseudoconvex Functions

A differentiable function f is pseudoconvex if

$$f(x'') < f(x') \Rightarrow \left[\nabla f(x')\right]^T (x'' - x') < 0$$

or equivalently,

$$\left[\nabla f(x')\right]^T (x'' - x') \geq 0 \Rightarrow f(x'') \geq f(x')$$

i.e., if the directional derivative is nonnegative in a direction d, then all points in that direction have a greater function value.

Objective function need only be pseudoconvex for stationarity to imply minimal solution!

quasiconvex but not convex

pseudoconvex but not convex

Sufficiency of KKT conditions

$$\begin{aligned} \text{Minimize } & f(x) \\ \text{s.t. } & g_i(x) \leq 0 \ \forall\, i = 1, \cdots m \\ & h_i(x) = 0 \ \forall\, i = 1, \cdots p \\ & x \in X \end{aligned}$$

Suppose that $(\hat{x}, \hat{\lambda}, \hat{\mu})$ satisfies the K-K-T conditions

$$- \nabla f(\hat{x}) = \sum_{i=1}^{m} \hat{\lambda}_i \nabla g_i(\hat{x}) + \sum_{i=1}^{p} \hat{\mu}_i \nabla h_i(\hat{x})$$

$$\hat{\lambda}_i\, g_i(\hat{x}) = 0 \ \forall\, i = 1, \cdots m$$

$$\hat{\lambda}_i \geq 0 \ \forall\, i = 1, \cdots m$$

$$\hat{x} \in X$$

K-K-T Sufficiency Conditions

Let $I = \{ i \mid g_i(\hat{x}) = 0 \}$ be the indices of the tight inequality constraints at $\hat{x}$

$J = \{ i \mid \hat{\mu}_i > 0 \}$

$K = \{ i \mid \hat{\mu}_i < 0 \}$

Suppose that
- f is pseudoconvex
- g_i is quasiconvex for $i \in I$
- h_i is quasiconvex for $i \in J$
- h_i is quasiconcave for $i \in K$

Then $\hat{x}$ is a globally optimal solution.