Optimization with (in)equality constraints

Evelien van der Hurk

DTU Management Engineering

Today's Topics in Optimization

- Part I: Equality and Inequality Constraints
- Part II:
    - Lagrange Multipliers
    - Lagrangian Relaxation
    - Lagrangian Duality

Last week: The Four-Step Method

The Four-Step Method for optimization

1. Model the problem and establish existence of global solutions (*e.g. Weierstrass*)
2. Write down the equation(s) of the first order necessary conditions (*e.g. Fermat*)
3. Investigate these equations
4. Write down the conclusion

Last week: Equality Constrained Optimization: The Lagrange multiplier rule

Lagrange multiplier rule

Given a problem $f_0(x) \to extr,\ f_i(x) = 0,,\ i \leq i \leq m$.
Assume that this problem is smooth at $\hat{x}$ in the following sense. A function $f_0 : U_n(\hat{x}, \epsilon) \to \mathbb{R}$ is differentiable at $\hat{x}$ and the functions $f_i : U_n(\hat{x}, \epsilon) \to \mathbb{R}, 1 \leq i \leq m$, are continuously differentiable at $\hat{x}$.
If $\hat{x}$ is a point of local extremum of this problem, then it is a stationary point of the Lagrange function of the problem for a suitable nonzero selection of Lagrange multipliers $\lambda \in (\mathbb{R}^{m+1})'$, that is:
$L_x(\hat{x}, \lambda) = 0_n^T \iff \sum_{i=0}^m \lambda_i f_i'(\hat{x}) = 0_n^T \iff \sum_{i=0}^m \lambda_i \frac{\delta f_i}{\delta x_j}(\hat{x}) = 0, 1 \leq j \leq n$

To remember it, you can write:

$\frac{\delta L}{\delta x_j} = 0, 1 \leq j \leq n, \frac{\delta L}{\delta \lambda_i} = 0, 1 \leq i \leq m$ Where you need to verify that indeed multiplier $\lambda_0 = 1$

Practice question:

$f_0(x) = x_1^2 + 12x_1x_2 + 2x_2^2 \to \text{extr},\ f_1(x) = 4x_1^2 + x_2^2 - 25 = 0$
Using the four-step method, solve the problem:
Solution:

1. Global extrema exist by Weierstrass (continuous function on closed interval)

Practice question:

$f_0(x) = x_1^2 + 12x_1x_2 + 2x_2^2 \to$ extr, $f_1(x) = 4x_1^2 + x_2^2 - 25 = 0$

Using the four-step method, solve the problem:

Solution:

1. Global extrema exist by Weierstrass (continuous function on closed interval)
2. Lagrange function: $L = \lambda_0(x_1^2 + 12x_1x_2 + 2x_2^2) + \lambda_1(4x_1^2 + x_2^2 - 25)$
   Lagrange: $L_x = 0_2^T \to$
   $\frac{\delta L}{\delta x_1} = \lambda_0 (2x_1 + 12x_2) + \lambda_1 (8x_1) = 0$ & $\frac{\delta L}{\delta x_2} = \lambda_0 (12x_1 + 4x_2) + \lambda_1 (2x_2) = 0$
   We put $\lambda_0 = 1$, as we may: if $\lambda_0 = 0$, then $\lambda_1 \neq 0$ and $x_1 = x_2 = 0$, contradicting the constraint.

Practice question:

$f_0(x) = x_1^2 + 12x_1x_2 + 2x_2^2 \to$ extr, $f_1(x) = 4x_1^2 + x_2^2 - 25 = 0$
Using the four-step method, solve the problem:

Solution:

1. Global extrema exist by Weierstrass (continuous function on closed interval)
2. Lagrange function: $L = \lambda_0(x_1^2 + 12x_1x_2 + 2x_2^2) + \lambda_1(4x_1^2 + x_2^2 - 25)$
   Lagrange: $L_x = 0_2^T \to$
   $\frac{\delta L}{\delta x_1} = \lambda_0 (2x_1 + 12x_2) + \lambda_1 (8x_1) = 0$ & $\frac{\delta L}{\delta x_2} = \lambda_0 (12x_1 + 4x_2) + \lambda_1 (2x_2) = 0$
   We put $\lambda_0 = 1$, as we may: if $\lambda_0 = 0$, then $\lambda_1 \neq 0$ and $x_1 = x_2 = 0$, contradicting the constraint.
3. Eliminate λ_1:
   $x_1x_2 + 6x_2^2 = 24x_1^2 + 8x_1x_2$
   This can be rewritten as: $6\left(\frac{x_2}{x_1}\right)^2 - 7\left(\frac{x_2}{x_1}\right) - 24 = 0$, provided $x_1 \neq 0$. This gives $x_2 = \frac{8}{3}x_1$ or $x_2 = -\frac{3}{2}x_1$. In the first (resp. second) case we get $x_1 = \pm\frac{3}{2}$ and so $x_2 = \pm 4$ (respectively $x_1 = \pm 2$, and so with reverse sign to x_1, $x_2 = \pm 3$), using the equality constraint. Compare: $f_0(2, -3) = f_0(-2, 3) = -50$ and $f_0(\frac{3}{2}, 4) = f_0(-\frac{3}{2}, -4) = 106\frac{1}{4}$.

Practice question:

$f_0(x) = x_1^2 + 12x_1x_2 + 2x_2^2 \rightarrow$ extr, $f_1(x) = 4x_1^2 + x_2^2 - 25 = 0$

Using the four-step method, solve the problem:

Solution:

1. Global extrema exist by Weierstrass (continuous function on closed interval)
2. Lagrange function: $L = \lambda_0(x_1^2 + 12x_1x_2 + 2x_2^2) + \lambda_1(4x_1^2 + x_2^2 - 25)$
   Lagrange: $L_x = 0_2^T \rightarrow$
   $\frac{\delta L}{\delta x_1} = \lambda_0\left(2x_1 + 12x_2\right) + \lambda_1\left(8x_1\right) = 0$ & $\frac{\delta L}{\delta x_2} = \lambda_0\left(12x_1 + 4x_2\right) + \lambda_1\left(2x_2\right) = 0$
   We put $\lambda_0 = 1$, as we may: if $\lambda_0 = 0$, then $\lambda_1 \neq 0$ and $x_1 = x_2 = 0$, contradicting the constraint.
3. Eliminate λ_1:
   $x_1x_2 + 6x_2^2 = 24x_1^2 + 8x_1x_2$
   This can be rewritten as: $6\left(\frac{x_2}{x_1}\right)^2 - 7\left(\frac{x_2}{x_1}\right) - 24 = 0$, provided $x_1 \neq 0$. This gives $x_2 = \frac{8}{3}x_1$ or $x_2 = -\frac{3}{2}x_1$. In the first (resp. second) case we get $x_1 = \pm\frac{3}{2}$ and so $x_2 = \pm 4$ (respectively $x_1 = \pm 2$, and so with reverse sign to x_1, $x_2 = \pm 3$), using the equality constraint. Compare: $f_0(2, -3) = f_0(-2, 3) = -50$ and $f_0(\frac{3}{2}, 4) = f_0(-\frac{3}{2}, -4) = 106\frac{1}{4}$.
4. $(2, -3)$ and $(-2, 3)$ are global minima and $(\frac{3}{2}, 4)$ and $-\frac{3}{2}, -4)$ global maxima

Equality, Inequality Constrained Optimization

Inequality Constraints

General Problem

$$\begin{aligned} \text{minimize} \quad & f_0(\boldsymbol{x}) \\ \text{subject to:} \quad & f_i(\boldsymbol{x}) \quad \leq 0 \qquad 1 \leq i \leq m \end{aligned}$$

- Lagrange function of this:

$$L(x, \lambda) = \sum_{i=0}^{m} \lambda_i f_i(x), \lambda = (\lambda_0, ..., \lambda_m) \tag{1}$$

- KKT conditions hold at $\hat{x}$ for a nonzero selection of λ with $\lambda_0 \geq 0$ when:
  - $L(\hat{x}, \lambda) = 0_n^T$ *The stationarity condition*
  - $\lambda_i \geq 0, 1 \leq i \leq m$ *The nonnegativity conditions*
  - $\lambda_i f_i(\hat{x}) = 0, 1 \leq i \leq m$ *The conditions of complimentary slackness*

Only essential difference Lagrange: $\lambda_i \geq 0$. Complimentary slackness: either $\lambda_i = 0$ or $f_i(\hat{x}) = 0$: 2^m cases

Constraint qualification

Slater point

A point $\bar{x}$ is a *Slater point* if

- $f_i(\boldsymbol{x}) < 0$ for $1 \leq i \leq m$

KKT Necessary

Assume that a problem with inequality constrainst as defined before is smooth at $\hat{x}$, and convex in the sense that functions $f_i : U_n(\hat{x}, \epsilon) \to \mathbb{R}, 0 \leq i \leq m$ are differentiable at $\hat{x}$, and they are convex. Then, if $\hat{x}$ is a point of minimum, there exists a nonzero selection of Lagrange mulipliers λ such that the KKT conditions hold at $\bar{x}$.

KKT Sufficiency

1. If the KKT conditions hold at $\hat{x}$ for a set of Lagrange multipliers λ with $\lambda_0 = 1$, then $\hat{x}$ is a point of minimum
2. Assume there exists a Slater point. If the KKT conditions hold for a selection of Lagrange multipliers λ, then λ_0 cannot be zero.

KKT remarks

KKT advantages

Both Lagrange and KKT leave to investigating as many as 2^m cases.

- Great advantage KKT: sufficiency. If we find a point, we do not gave to check the remaining cases, as the KKt provide a "certificate of optimality".
- Small advantage: Only have to consider non negative Lagrange multipliers

Example

Problem

Model the problem to find the point closest to (2,3) with the sum of the coordinates not larger than 2 and with the first coordinate not larger than 2 in absolute value

Example

Problem

Model the problem to find the point closest to (2,3) with the sum of the coordinates not larger than 2 and with the first coordinate not larger than 2 in absolute value

$$f_0(x) = (x_1 - 2)^2 + (x_2 - 3)^2 \to \min$$
$$f_1(x) = x_1 + x_2 - 2 \leq 0, f_2(x) = x_1^2 - 4 \leq 0$$

Example

Problem

Model the problem to find the point closest to (2,3) with the sum of the coordinates not larger than 2 and with the first coordinate not larger than 2 in absolute value

$$f_0(x) = (x_1 - 2)^2 + (x_2 - 3)^2 \to \min$$
$$f_1(x) = x_1 + x_2 - 2 \leq 0, f_2(x) = x_1^2 - 4 \leq 0$$

1. As the problem is convex, we do not need to apply Weierstrass
2. Lagrange function:$L(x, \lambda) = \lambda_0((x_1 - 2)^2 + (x_2 - 3)^2) + \lambda_1(x_1 + x_2 - 2) + \lambda_2(x_1^2 - 4)$ put $\lambda_0 = 1, (0, 0)$ is a Slater point.

Example - cont. Formulate KKT conditions and evaluate cases

$$
\begin{aligned}
\frac{\delta L}{\delta x_1} = 2(x_1 - 2) + \lambda_1 + 2\lambda_2 x_1 = 0 \\
\frac{\delta L}{\delta x_2} = 2(x_2 - 3) + \lambda_1 = 0 \\
\lambda_1, \lambda_2 \geq 0 \\
\lambda_1(x_1 + x_2 - 2) = 0 \\
\lambda_2(x_1^2 - 4) = 0
\end{aligned}
$$

Example - cont. Formulate KKT conditions and evaluate cases

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

$$\frac{\delta L}{\delta x_2} = 2(x_2 - 3) + \lambda_1 = 0$$

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

$$\lambda_1(x_1 + x_2 - 2) = 0$$

$$\lambda_2(x_1^2 - 4) = 0$$

Distinguish cases:

Case 1: no constraint is tight:

$\lambda_1 = \lambda_2 = 0, 2(x_1 - 2) = 0, 2(x_2 - 3) = 0 \rightarrow (2, 3)$, however, this point is not admissible.

Example - cont. Formulate KKT conditions and evaluate cases

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

$$\frac{\delta L}{\delta x_2} = 2(x_2 - 3) + \lambda_1 = 0$$

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

$$\lambda_1(x_1 + x_2 - 2) = 0$$

$$\lambda_2(x_1^2 - 4) = 0$$

Distinguish cases:

Case 1: no constraint is tight:
$\lambda_1 = \lambda_2 = 0, 2(x_1 - 2) = 0, 2(x_2 - 3) = 0 \rightarrow (2, 3)$, however, this point is not admissible.

Case 2: Only the second constraint is tight. Then KKT:
$\lambda_1 = 0, x_1^2 = 4, 2(x_2 - 3) = 0$, $2(x_1 - 2) + 2\lambda_2 x_1 = 0, \lambda_2 \geq 0$. However, the resulting point $(2, 3)$ is not admissible.

Example - cont. Formulate KKT conditions and evaluate cases

$$\frac{\delta L}{\delta x_1} = 2(x_1 - 2) + \lambda_1 + 2\lambda_2 x_1 = 0$$
$$\frac{\delta L}{\delta x_2} = 2(x_2 - 3) + \lambda_1 = 0$$
$$\lambda_1, \lambda_2 \geq 0$$
$$\lambda_1(x_1 + x_2 - 2) = 0$$
$$\lambda_2(x_1^2 - 4) = 0$$

Distinguish cases:
Case 3: Only the first constraint is tight. Then KKT:
$\lambda_2 = 0, x_2 + x_2 = 2, 2(x_1 - 2) + \lambda_1 = 0$
$2(x_2 - 3) + \lambda_1 = 0, \lambda_1 \geq 0$. The resulting point $(\frac{1}{2}, \frac{3}{2}$ with $\lambda_1 = 3$ is admissible, and therefore it follows that it is a global solution.