CE 191: Civil and Environmental Engineering Systems Analysis

LEC 13 : Lagrange Multipliers & KKT Conditions

Professor Scott Moura
Civil & Environmental Engineering
University of California, Berkeley

Fall 2014

Constrained Nonlinear Programming

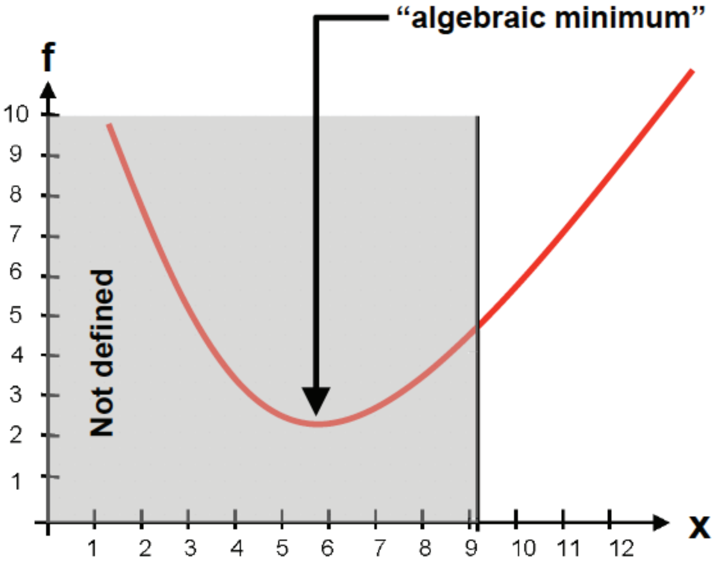
Abstract Optimization Problem

$$\begin{aligned} \min \quad & f(x) & \\ \text{s. to} \quad & g_i(x) \leq 0, & i = 1, \cdots, m \\ & h_j(x) = 0, & j = 1, \cdots, l \end{aligned}$$

Questions / Issues

1. What, exactly, is the definition of a minimum? local and global
2. Does a solution even exist? feasibility
3. Is it unique? if it's convex, then yes
4. What are the necessary & sufficient conditions to be a solution?
5. How do we solve?

Simple Graphical Example



Method of Lagrange Multipliers

Equality Constrained Optimization Problem

$$\begin{aligned} \min \quad & f(x) \\ \text{s. to} \quad & h_j(x) = 0, \qquad j = 1, \cdots, l \end{aligned}$$

Lagrangian

Introduce the so-called “Lagrange multipliers” $\lambda_j, j = 1, \cdots, l$, i.e. one for each equality constraint. The Lagrangian is

$$\begin{aligned} L(x) &= f(x) + \sum_{j=1}^{l} \lambda_j h_j(x) \\ &= f(x) + \lambda^T h(x) \end{aligned}$$

First order Necessary Condition (FONC)

If a local minimum x^* exists, then it satisfies

$$\nabla L(x^*) = \nabla f(x^*) + \lambda^T \nabla h(x^*) = 0$$

Remarks

Remark 1 - Only necessary

This condition is only necessary. That is, if a local minimum x^* exists, then it must satisfy the FONC. However, a design x which satisfies the FONC isn't necessarily a local minimum.

Remark 2 - Convexity $\Rightarrow$ Necessary and sufficient

If the optimization problem is convex, then the FONC is necessary and sufficient. That is, a design x which satisfies the FONC is also a local minimum.

Example 1: Equality Constrained QP

$$\begin{array}{rl} \min & \frac{1}{2}x^T Qx + Rx \\ \text{s. to} & Ax = b \end{array}$$

Form the Lagrangian:

$$L(x) = \frac{1}{2}x^T Qx + Rx + \lambda^T (Ax - b)$$

Then the FONC is:

$$\nabla_x L(x^*) = Qx^* + R + A^T\lambda = 0$$

Combining the FONC with the equality constraint:

$$\begin{bmatrix} Q & A^T \\ A & 0 \end{bmatrix} \begin{bmatrix} x^* \\ \lambda \end{bmatrix} = \begin{bmatrix} -R \\ b \end{bmatrix}$$

provides a set of linear equations, which can be solved directly!

Example 2: Circle and plane

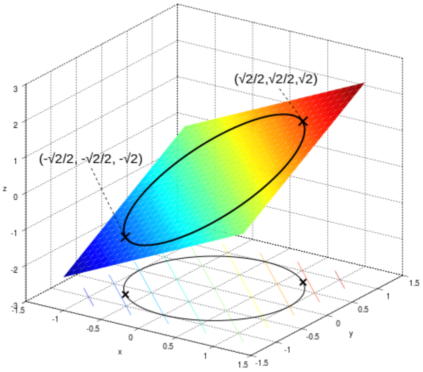
$$\begin{array}{ll} \min & f(x,y) = x + y \\ \text{s. to} & x^2 + y^2 = 1 \end{array}$$

Form the Lagrangian:

$$L(x, y, \lambda) = x + y + \lambda(x^2 + y^2 - 1)$$

The FONC are

$$\begin{aligned} \frac{\partial L}{\partial x} &= 1 + 2\lambda x = 0 \\ \frac{\partial L}{\partial y} &= 1 + 2\lambda y = 0 \\ \frac{\partial L}{\partial \lambda} &= x^2 + y^2 - 1 = 0 \end{aligned}$$



$$\begin{aligned} (x^*, y^*) &= \left(\pm\frac{\sqrt{2}}{2}, \pm\frac{\sqrt{2}}{2} \right) \\ f(x^*, y^*) &= \pm\sqrt{2} \\ \lambda &= \mp 1/\sqrt{2} \end{aligned}$$

Karush-Kuhn-Tucker (KKT) Conditions

General Constrained Optimization Problem

$$
\begin{aligned}
\min \quad & f(x) \\
\text{s. to} \quad & g_i(x) \leq 0, \qquad i = 1, \cdots, m \\
& h_j(x) = 0, \qquad j = 1, \cdots, l
\end{aligned}
$$

If x^* is a local minimum, then the following necessary conditions hold:

$$\nabla f(x^*) + \sum_{i=1}^{m} \mu_i \nabla g_i(x^*) + \sum_{j=1}^{l} \lambda_j \nabla h_j(x^*) = 0, \qquad \text{Stationarity} \qquad (1)$$

$$
\begin{aligned}
g_i(x^*) &\leq 0, \quad i = 1, \cdots, m & \text{Feasibility} \qquad (2) \\
h_j(x^*) &= 0, \quad j = 1, \cdots, l & \text{Feasibility} \qquad (3) \\
\mu_i &\geq 0, \quad i = 1, \cdots, m & \text{Non-negativity} \qquad (4) \\
\mu_i g_i(x^*) &= 0, \quad i = 1, \cdots, m & \text{Complementary slackness} \qquad (5)
\end{aligned}
$$

Karush-Kuhn-Tucker (KKT) Conditions

General Constrained Optimization Problem

$$\begin{array}{rll} \min & f(x) & \\ \text{s. to} & g_i(x) \leq 0, & i = 1, \cdots, m \\ & h_j(x) = 0, & j = 1, \cdots, l \end{array}$$

If x^* is a local minimum, then the following necessary conditions hold: [Matrix form]

$$\nabla f(x^*) + \mu^T \nabla g(x^*) + \lambda^T \nabla h(x^*) = 0, \qquad \text{Stationarity} \tag{1}$$

$$g(x^*) \leq 0, \qquad \text{Feasibility} \tag{2}$$

$$h(x^*) = 0, \qquad \text{Feasibility} \tag{3}$$

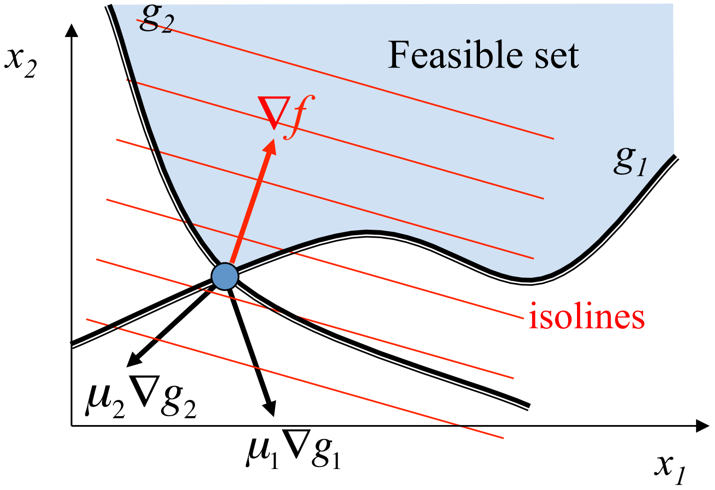
$$\mu \geq 0, \qquad \text{Non-negativity} \tag{4}$$

$$\mu^T g(x^*) = 0, \qquad \text{Complementary slackness} \tag{5}$$

Remarks

- Non-zero μ_i indicates $g_i \leq 0$ is active (true with equality).
- Conditions are necessary, only.
- If problem is convex, then the conditions are necessary and sufficient.
- Lagrange multipliers λ, μ are sensitivity to perturbations in constraints
  - In economics, this is called the “shadow price”
  - In control theory, this is called the “co-state”

Geometric Interpretation



Weighted sum of gradient vectors balances to zero

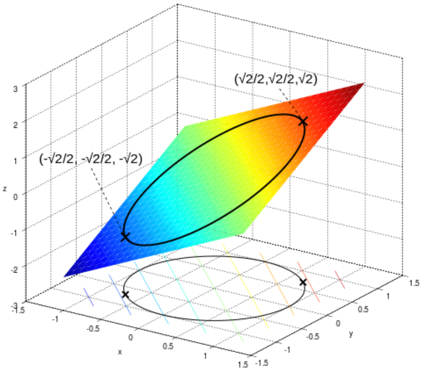
$$\nabla f(x^*) + \mu_1 \nabla g_1(x^*) + \mu_2 \nabla g_2(x^*) = 0$$

Example 3: Circle and plane REDUX

$$\begin{array}{ll} \min & f(x,y) = x + y \\ \text{s. to} & x^2 + y^2 \leq 1 \end{array}$$

The KKT conditions are

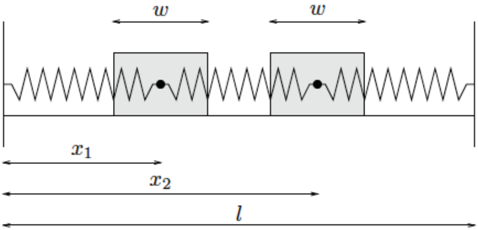
$$\begin{aligned} \frac{\partial L}{\partial x} &= 1 + 2\mu x = 0 \\ \frac{\partial L}{\partial y} &= 1 + 2\mu y = 0 \\ \frac{\partial L}{\partial \lambda} &= x^2 + y^2 - 1 \leq 0 \\ & \mu \geq 0 \\ & \mu\left(x^2 + y^2 - 1\right) = 0 \end{aligned}$$



$$\begin{aligned} (x^*, y^*) &= \left(-\frac{\sqrt{2}}{2}, -\frac{\sqrt{2}}{2}\right) \\ f(x^*, y^*) &= -\sqrt{2} \\ \mu &= 1/\sqrt{2} \end{aligned}$$

Ex 4: Mechanics (Physics 7A) Interpretation

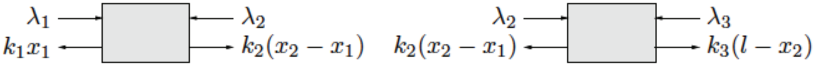
Find equilibrium,
i.e. minimize potential energy
subject to kinematic constraints



$$
\begin{aligned}
\min \quad & f(x_1, x_2) = \frac{1}{2} k_1 x_1^2 + \frac{1}{2} k_2 (x_2 - x_1)^2 + \frac{1}{2} k_3 (l - x_2)^2 \\
\text{s. to} \quad & x_1 - \frac{w}{2} \geq 0, \\
& x_1 + \frac{w}{2} \leq x_2 - \frac{w}{2}, \\
& x_2 + \frac{w}{2} \leq l
\end{aligned}
$$

This is a QP

Ex 4: Mechanics (Physics 7A) Interpretation



With $\lambda_1, \lambda_2, \lambda_3$ as the Lagrange multipliers, the KKT conditions are:

$\lambda_i \geq 0$ for non-negativity,

$$\lambda_1 \left(\frac{w}{2} - x_1 \right) = 0, \qquad \lambda_2 \left(x_1 - x_2 + w \right) = 0, \qquad \lambda_3 \left(x_2 + \frac{w}{2} - l \right) = 0$$

for complementary slackness, and

$$\begin{bmatrix} k_1 x_1 - k_2 (x_2 - x_1) \\ k_2 (x_2 - x_1) - k_3 (l - x_2) \end{bmatrix} + \lambda_1 \begin{bmatrix} -1 \\ 0 \end{bmatrix} + \lambda_2 \begin{bmatrix} 1 \\ -1 \end{bmatrix} + \lambda_3 \begin{bmatrix} 0 \\ 1 \end{bmatrix} = 0$$

for stationarity.

Interestingly, the λ_i's can be interpreted as contact forces.

Additional Reading

Boyd & Vandenberghe, Section 5.5

Papalambros & Wilde, Chapter 5