From Lagrangian Multipliers to KKT Conditions

A Geometric Journey Through Constrained Optimization

Zirui Zhang

Cheng Kar-Shun Robotics Institute
The Hong Kong University of Science and Technology

September 27, 2025

Optimization with Equality Constraints

Consider an optimization problem with equality constraints:

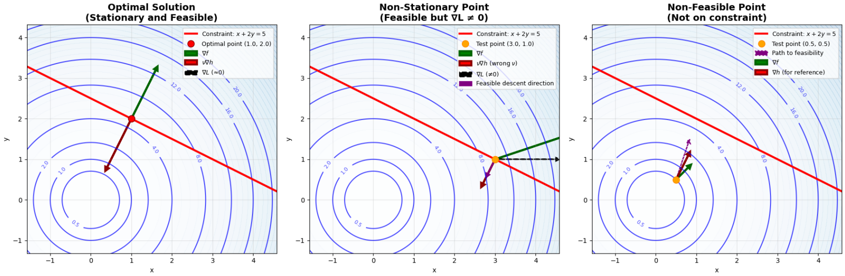
$$\min_x f(x) \quad \text{subject to} \quad h_j(x) = 0$$

how to characterize the necessary conditions for the optimal solution x^* ?

Geometric Insight

Consider a quadratic objective function with one linear equality constraint:

$$\min_{x,y} x^2 + y^2 \quad \text{subject to} \quad x + 2y = 5$$

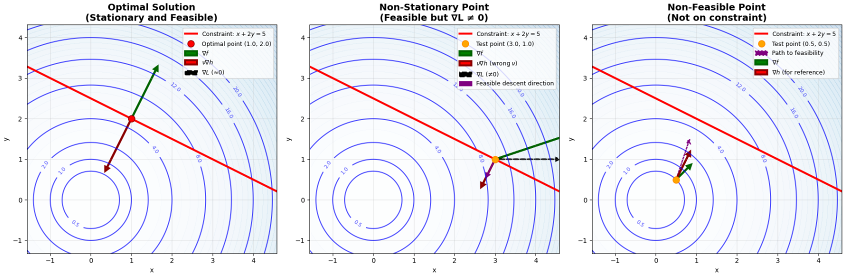


$\nabla f(x^*)$ must lie within the linear subspace spanned by $\{\nabla h_j(x^*)\}$; otherwise, the function value could be further decreased by moving along a feasible direction.

Geometric Insight (cont.)

$\nabla f(x^*)$ must lie in the space spanned by $\{\nabla h_j(x^*)\}$. This means there exist multipliers $\{\nu_j^*\}$ such that:

$$\nabla f(x^*) + \sum_j \nu_j^* \nabla h_j(x^*) = 0$$



Lagrangian Function and Optimality Conditions

We introduce the **Lagrangian function** as a tool to characterize optimality:

$$\mathscr{L}(x, v) = f(x) + \sum_j v_j h_j(x)$$

The necessary conditions for optimality can be expressed as stationarity of the Lagrangian:

$$\nabla \mathscr{L}(x^*, v^*) = 0 \quad \Longleftrightarrow \quad \begin{cases} \nabla_x \mathscr{L} = \nabla f(x^*) + \sum_j v_j^* \nabla h_j(x^*) = 0 \\ \nabla_v \mathscr{L} = [\dots, h_j(x^*), \dots]^\top = 0 \end{cases}$$

A Min-Max Interpretation

The Lagrangian function also leads to a powerful dual interpretation:

$$\max_{\nu} \mathscr{L}(x,\nu) = \begin{cases} f(x), & h_j(x) = 0 \\ \infty, & \text{otherwise} \end{cases}$$

The original constrained problem is equivalent to the following **min-max problem**:

$$\min_x f(x), \text{ s.t. } h_j(x) = 0 \iff \min_x \max_{\nu} \mathscr{L}(x,\nu)$$

The Dual Problem and Weak Duality

Solution for $\min_x \max_\nu \mathscr{L}(x,\nu)$ may be non-continuous, but solution for $\max_\nu \min_x \mathscr{L}(x,\nu)$ is easy if $\mathscr{L}(x,\nu)$ is **tractable**. We can form the dual problem by swapping the order of the min and the max:

$$\max_\nu d(\nu) = \max_\nu \min_x \mathscr{L}(x,\nu) \leq \min_x \max_\nu \mathscr{L}(x,\nu)$$

Under some conditions, equality holds, which means **strong duality** holds.

Uzawa's Method (Dual Ascent)

The gradient of the dual function can be computed as:

$$\nabla d(v) = h[x^*(v)] \quad \text{where} \quad x^*(v) = \arg\min_x \mathscr{L}(x, v)$$

This leads to Uzawa's Method:

1. **Minimization** (x-step):
$$x^{k+1} = \arg\min_x \mathscr{L}(x, v^k)$$
2. **Ascent** (v-step):
$$v^{k+1} = v^k + \alpha^k h(x^{k+1})$$

   where $\alpha^k > 0$ is the step size.

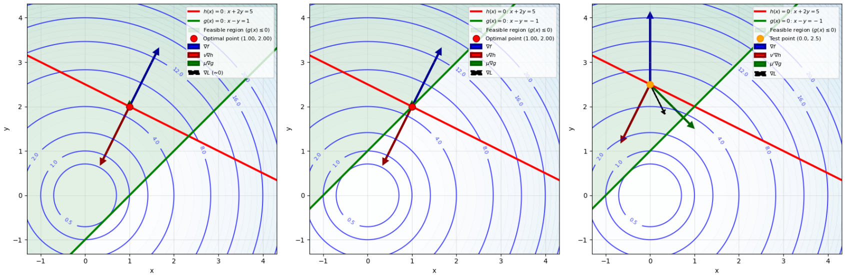
General Constrained Optimization

Consider a general optimization problem with equality and inequality constraints:

$$\begin{aligned} &\min_x f(x) \\ \text{s.t. } &g_i(x) \leq 0 \\ &h_j(x) = 0 \end{aligned}$$

The question does not change: how to characterize the necessary conditions for the optimal solution x^* ?

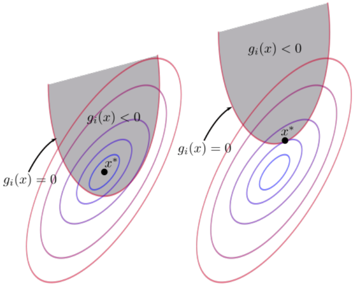
Geometric Insight



Challenge:

1. **Directionality:** On the boundary, ∇g_i points towards the exterior of the feasible region. To prevent f from pushing the point into an infeasible area, ∇f must have a component opposite to $\nabla g_i \implies \mu_i \geq 0$.
2. **Activity Identification:** The optimum may lie in the interior of the region with $g_i < 0$ or on the boundary with $g_i = 0 \implies \mu_i g_i = 0$.

Summary



1. **Stationarity:** $0 \in \partial_x[f(x) + \sum_i \mu_i g_i(x) + \sum_j \nu_j h_j(x)]$
2. **Complementary Slackness:** $\mu_i g_i(x) = 0$
3. **Primal Feasibility:** $g_i(x) \leq 0, h_j(x) = 0$
4. **Dual Feasibility:** $\mu_i \geq 0$

Thank you for listening !

Zirui Zhang