Constrained Optimization

Prithwijit Guha
Dept. of EEE, IIT Guwahati

Overview

- The Constrained Optimization Problem
- Contour Plot and Gradients
- Objective Function and Gradients
- The KKT Conditions
  - Necessary Conditions
  - Sufficient Conditions
- The Primal-Dual Formulation

Constrained Optimization: Definition & Formulation

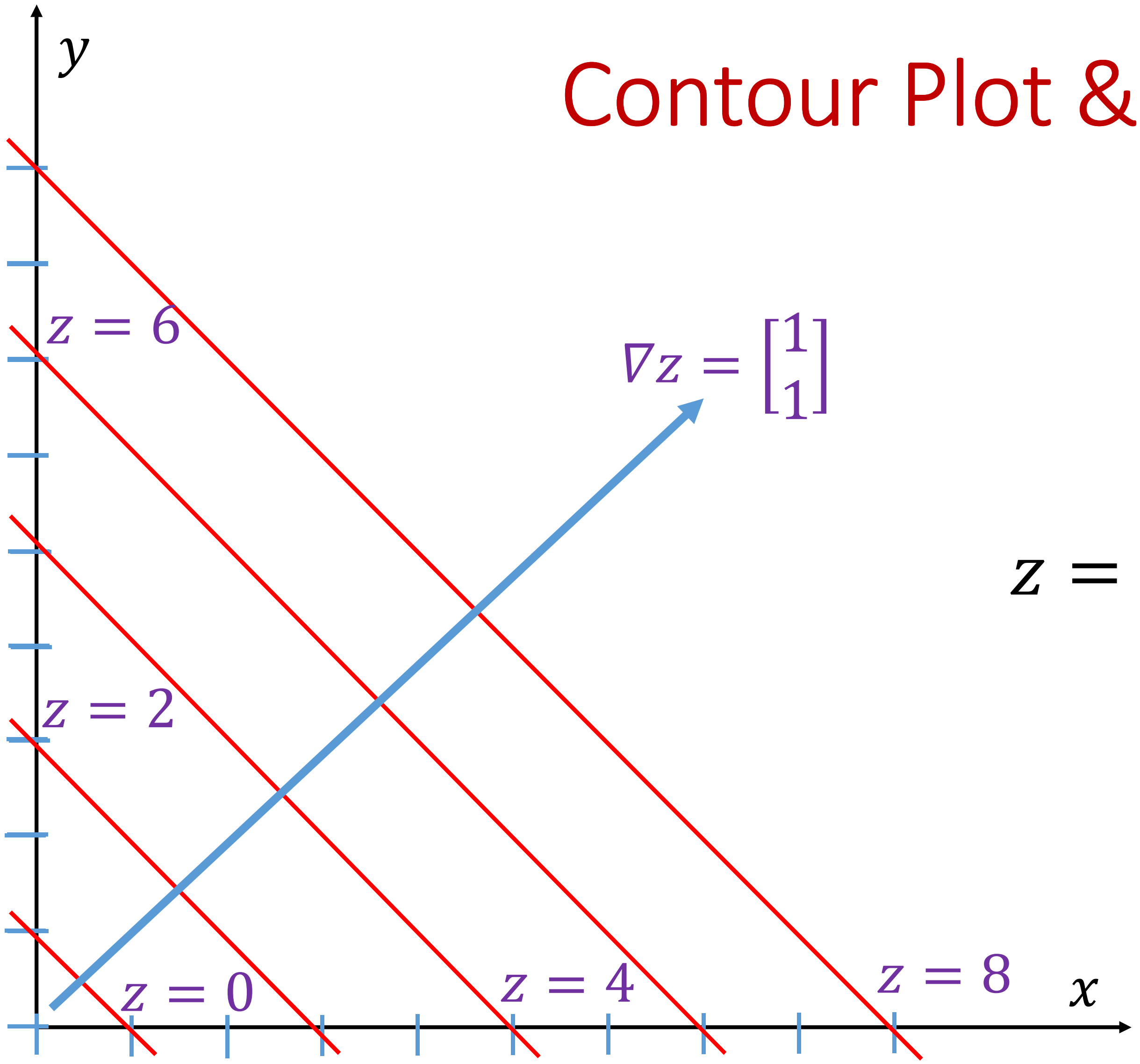
Constrained Optimization is the Process of Maximizing or Minimizing a Desired Objective Function while Satisfying a Set of Prevailing Constraints

$$Minimize\ f(X)$$

Subject To

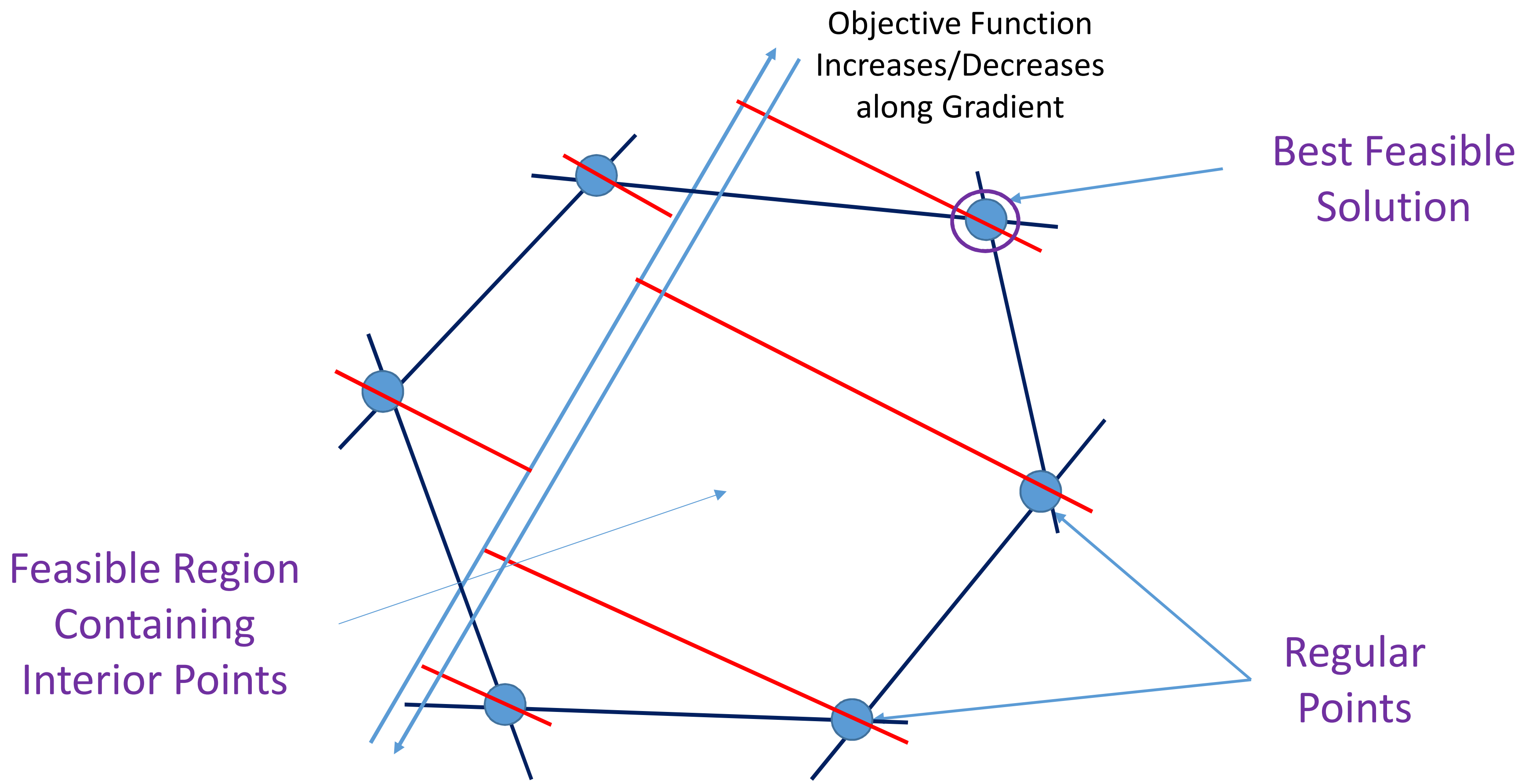
$$g_i(X) \leq 0; i = 1, \dots m \qquad\qquad h_j(X) = 0; j = 1, \dots l$$

Contour Plot & Gradient

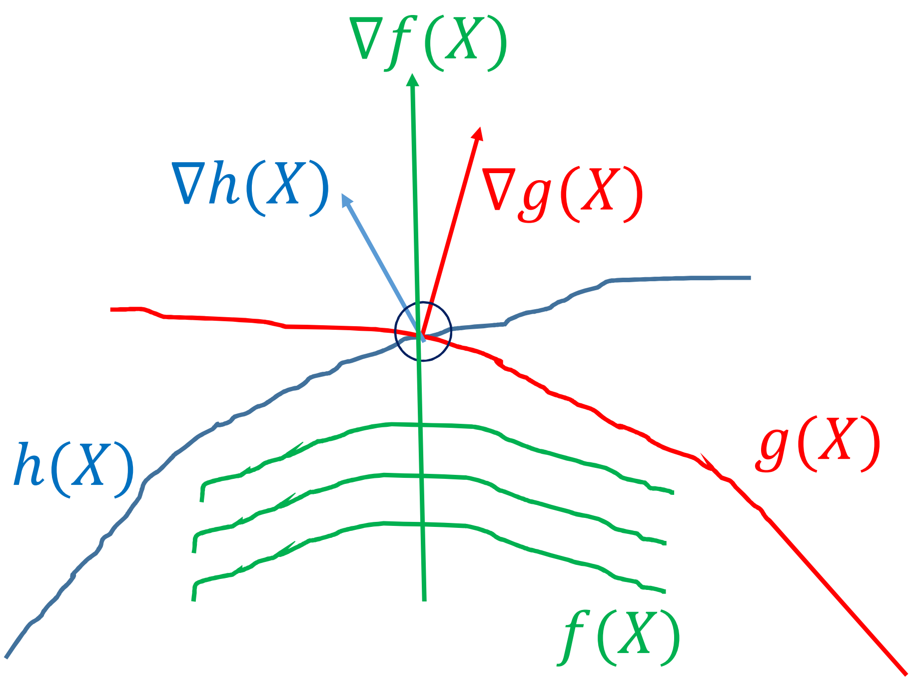


$$z = f(x, y) = x + y - 1$$

Objective Functions & Gradients: Linear Case



Objective Functions & Gradients (Contd.)



$$Minmize\ f(X)$$

$$Subject\ to \quad g(X) \le 0$$

$$h(X) = 0$$

Observation:
$\nabla f(X), \nabla g(X)\ and\ \nabla h(X)$
pass through the same point

Thus, we can write
$$\nabla f(X) + \mu \nabla g(X) + \lambda \nabla h(X) = 0$$

Karush-Kuhn-Tucker (KKT) Conditions

Minimize $f(X)$ subject to $g_i(X) \leq 0; i = 1 \dots m$ and $h_j(X) = 0; j = 1 \dots l$

Construct Lagrangian Function ***L***

$$L(X; \boldsymbol{\mu}, \boldsymbol{\lambda}) = f(X) + \sum_{i=1}^{m} \mu_i g_i(X) + \sum_{j=1}^{l} \lambda_j h_j(X)$$

Where, $\boldsymbol{\mu} = \{ \mu_1 \dots \mu_i \dots \mu_m\}$ and $\boldsymbol{\lambda} = \{\lambda_1 \dots \lambda_j \dots \lambda_l\}$

KKT Conditions (Contd.)

$$\textbf{Optimality}: \nabla_X L(X; \boldsymbol{\mu}, \boldsymbol{\lambda}) = \nabla_X f(X) + \sum_{i=1}^{m} \mu_i \nabla_X g_i(X) + \sum_{j=1}^{l} \lambda_j \nabla_X h_j(X) = 0$$

$\textbf{Non} - \textbf{negativity}: \mu_i \geq 0\ ; i = 1 \ldots m$

$\textbf{Complementarity}: \mu_i g_i(X) = 0; i = 1, \ldots m$

$\textbf{Feasibility}: g_i(X) \leq 0; i = 1, \ldots m$ and $h_i(X) = 0; j = 1, \ldots l$

Necessary Conditions

KKT Conditions (Contd.)

X^{*}: Optimal Solution of $\nabla_X L(X; \boldsymbol{\mu}, \boldsymbol{\lambda}) = 0$

Hessian Matrix $\nabla_X^2 L(X^{*}; \mu, \lambda)$ *is Positive Definite in* $\boldsymbol{T} \subset \mathbb{R}^n$

$\boldsymbol{T} = \left\{y: [\nabla g_i(X^{*})^T y = 0; i = 1 \ldots m] \wedge \left[\nabla h_j(X^{*})^{\mathrm{T}} \mathrm{y} = 0; \mathrm{j} = 1 \ldots l\right]\right\}$

Sufficient Condition: $\boldsymbol{y}^{\boldsymbol{T}} \nabla_X^2 L(X^{*}; \mu, \lambda)\mathrm{y} > 0; \|y\|_2 \neq 0; y \in \boldsymbol{T}$

Stronger Condition: $\boldsymbol{y}^{\boldsymbol{T}} \nabla_X^2 L(X^{*}; \mu, \lambda)\mathrm{y} > 0; \|y\|_2 \neq 0; y \in \mathbb{R}^{\boldsymbol{n}}$

Problem

$$\text{Minimize } f(x_1, x_2) = 2x_1 + x_2$$

Subject To

$$h(x_1, x_2) = x_1^2 + x_2^2 - 1 = 0$$

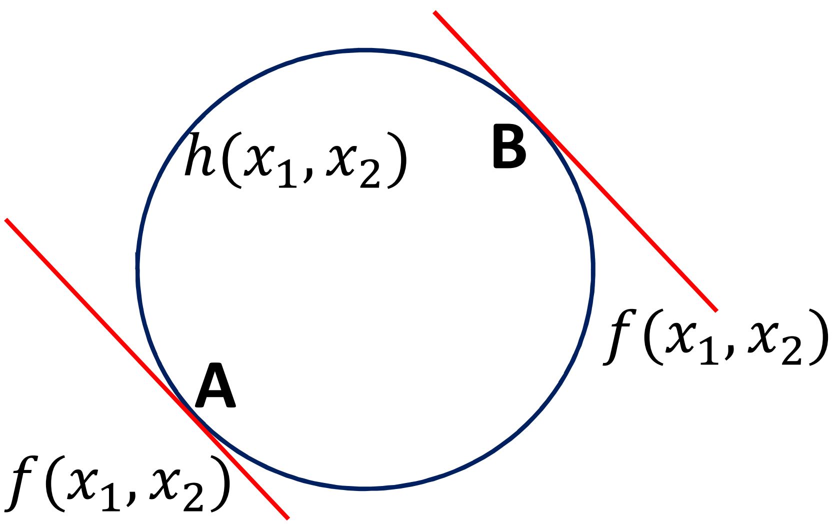
Solution

$$L(x_1, x_2; \lambda) = 2x_1 + x_2 + \lambda(x_1^2 + x_2^2 - 1)$$

$$\frac{\partial L}{\partial x_1} = 0 \rightarrow x_1 = -\frac{1}{\lambda}$$

$$\frac{\partial L}{\partial x_2} = 0 \rightarrow x_2 = -\frac{1}{2\lambda}$$

$$\lambda = \pm\frac{\sqrt{5}}{2}$$



$$A = \left(x_1 = -\frac{2}{\sqrt{5}}, x_2 = -\frac{1}{\sqrt{5}}\right)$$

$$B = \left(x_1 = \frac{2}{\sqrt{5}}, x_2 = \frac{1}{\sqrt{5}}\right)$$

Primal-Dual Formulation

Minimize $f(X)$ subject to $g_i(X) \leq 0; i = 1 \ldots m$ and $h_j(X) = 0; j = 1 \ldots l$

$X^*, \boldsymbol{\mu}^*, \boldsymbol{\lambda}^*$: Optimal Solution of the **Primal Problem** $\nabla_X L(X; \boldsymbol{\mu}, \boldsymbol{\lambda}) = 0$

$(X^*, \boldsymbol{\mu}^*, \boldsymbol{\lambda}^*)$ is a **Saddle Point** of $L(X; \boldsymbol{\mu}, \boldsymbol{\lambda})$ *i.e.*

$L(X^*, \boldsymbol{\mu}, \boldsymbol{\lambda}) \leq L(X^*, \boldsymbol{\mu}^*, \boldsymbol{\lambda}^*) \leq L(X, \boldsymbol{\mu}^*, \boldsymbol{\lambda}^*)$

$(X^*, \boldsymbol{\mu}^*, \boldsymbol{\lambda}^*)$ **solves the following dual problem**

$$Maximize\ L = f + \boldsymbol{\mu}^T \boldsymbol{g} + \boldsymbol{\lambda}^T \boldsymbol{h}$$
$$Subject\ To\ [\boldsymbol{\nabla} L = 0] \wedge [\boldsymbol{\mu} \geq 0]$$

Problem

$$\text{Minimize } f(x_1, x_2) = (x_1 - 3)^2 + (x_2 - 3)^2$$

Subject To

$$g_1(x_1, x_2) = 2x_1 + x_2 - 2 \leq 0$$

$$g_2(x_1, x_2) = -x_1 \leq 0$$

$$g_3(x_1, x_2) = -x_2 \leq 0$$

Solution

$$L(x_1, x_2; \boldsymbol{\mu}) = (x_1 - 3)^2 + (x_2 - 3)^2 + \mu_1(2x_1 + x_2 - 2) - \mu_2 x_1 - \mu_3 x_2$$

$$\frac{\partial L}{\partial x_1} = 0 \rightarrow x_1 = \frac{\mu_2}{2} - \mu_1 + 3$$

$$\frac{\partial L}{\partial x_2} = 0 \rightarrow x_2 = \frac{\mu_3}{2} - \frac{\mu_1}{2} + 3$$

$$Q(\boldsymbol{\mu}) = -\frac{5\mu_1^2}{4} - \frac{\mu_2^2}{4} - \frac{\mu_3^2}{4} + \frac{\mu_1\mu_3}{2} + \mu_1\mu_2 + 7\mu_1 - 3\mu_2 - 3\mu_3$$

Dual Function to be Maximized $\mathrm{w.r.t.}\ \boldsymbol{\mu}$

Applications

- Principal Component Analysis
- Linear Discriminant Analysis
- Factor Analysis
- Regularization
- Neural Network Learning
- Support Vector Machines
- Parameter Estimation And Many Others…

Summary

- Constrained Minimization
- Visualizations
  - Contour Plots & Gradients
  - Objective Functions & Gradients
  - Feasible Regions and Best Solutions
- Karush-Kuhn-Tucker Conditions
  - Necessary Conditions
  - Sufficient Conditions
- Primal-Dual Formulations
- Applications

Thank You