Module – 2.4

OPTIMIZATION IN FINITE DIMENSIONAL VECTOR SPACES:

AN OVERVIEW

S. Lakshmivarahan

School of Computer Science

University of Oklahoma

Norman, Ok – 73069, USA

varahan@ou.edu

A CLASSIFICATION – MINIMUN AND MAXIMUM

- Let f: R^n -> R and f ∈ $C^2(R^n)$

- f is twice continuously differentiable functional

$$\min_{x \in R^n} f(x) = \max_{x \in R^n} \{-f(x)\}$$

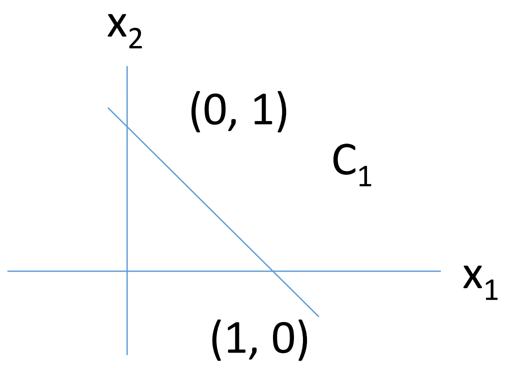
- Consider only minimization

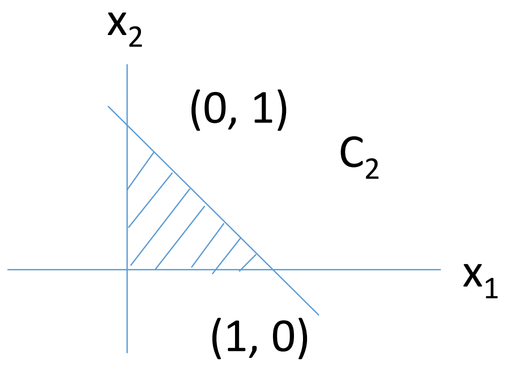
A CLASSIFICATION – UNI VS MULTI MODAL

- $N_\varepsilon(x) = \{ y \in R^n \mid \|y - x\| \leq \varepsilon \} \subset R^n$ called ε – neighborhood
- If $x^* \in R^n$ is such that $f(x^*) \leq f(y)$ for all $y \in N_\varepsilon(x^*)$, then x^* is a local minimum
- If $x^* \in R^n$ is such that $f(x^*) \leq f(y)$ for all y, then x^* is a global minimum
- A function that has a unique minimum is a unimodal function
- Otherwise, it is a multimodal function
- $f(x) = x(x^2 – 1)$ is multimodal, $f(x) = x^2$ is unimodal

A CLASSIFICATION – CONSTRAINED VS UNCONSTRAINED

- Let $C \subset R^n$ defined by a set of equations or inequalities
- $C_1 = \{ x \in R^2 \mid x_1 + x_2 = 1 \}$
- $C_2 = \{ x \in R^2 \mid x_1 \geq 0, x_2 \geq 0, \; x_1 + x_2 \leq 1 \}$
- Let $f(x) = x_1^2 + x_2^2$
- $\underset{x \in R^2}{\text{Min}} \, f(x)$ is an unconstrained minimization
- $\underset{x \in C_1}{\text{Min}} \, f(x)$ is a constrained minimization with equality constraints
- $\underset{x \in C_2}{\text{Min}} \, f(x)$ is a constrained minimization with inequality constraints
- Linear and non linear programing deal with minimization under inequality constraints
- In this course we will deal with unconstrained and equality constrained minimization problem only





A CLASSIFICATION – UNI VS MULTI-OBJECTIVE OPTIMIZATION

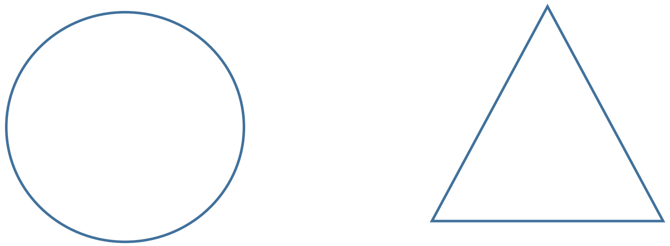
- If f: R^n -> R is the only function to be minimized, it is known as uni-objective minimization
- If f: R^n -> R^m with $f(x) = (f_1(x), f_2(x), \ldots f_m(x))^T$ where we want to minimize some and maximize others, it is called multi-objective optimization
- Automobile design – Maximize fuel efficiency, minimize cost, maximize safety and comfort is an example of multi-objective optimization
- In this course we only deal with uni-objective minimization

ROLE OF CONVEXITY IN MINIMIZATION

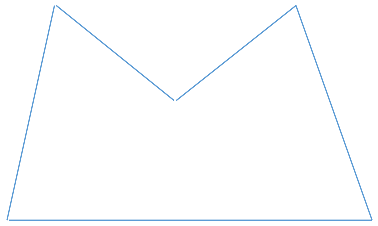
- Let S be a subset of R^n

- S is called a <u>convex set</u> if for every pair of points x and y in S, the points along the line segment joining x and y are also in S

$$\alpha x + (1 - \alpha)y \in S \text{ if } x, y \in S$$



Convex sets



Non-Convex set

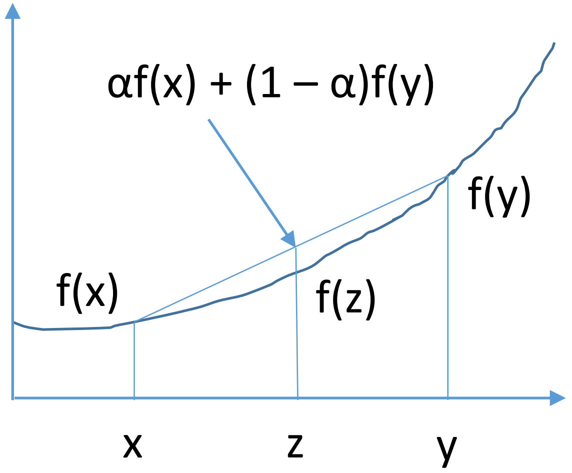
ROLE OF CONVEXITY IN MINIMIZATION

- Let S be a convex set in R^n and let x, y ∈ S
- A function f: S -> R is said to be a <u>convex function</u> if

$$f(\alpha x + (1 - \alpha)y) \leq \alpha f(x) + (1 - \alpha)f(y)$$

$$\text{for all } \alpha \in [0, 1]$$

- A convex function lies below the chord
- Let $z = \alpha x + (1 - \alpha)y$
- If f(x) is convex, -f(x) is concave
- $f(x) = x^2$ is convex but $g(x) = x^3$ is not



CHARACTERIZATION OF CONVEXITY

- If f ∈ $C^1(S)$ be a continuously differential function defined on a convex set S
- f is convex if and only if, for x, y ∈ S

$$f(y) \geq f(x) + (y - x)^T \nabla_x f(x)$$ -> curve lies above the tangent

- f(x) is strictly convex if strict inequality holds
- If f ∈ $C^2(S)$ be twice continuously differentiable function on a convex set S
- Then f is convex if and only if the Hessian $\nabla_x^2 f(x)$ is positive semi-definite. f is strictly convex if $\nabla_x^2 f(x)$ is positive definite

CONVEXITY AND UNIMODALITY

- f: S -> R and S is a convex set

- Then f has a unique minimum

- If $f \in C^2(S)$, then at this minimum $\nabla_x f(x) = 0$ and $\nabla_x^2 f(x)$ is positive definite

- $f = x^T A x - b^T x$ is a typical convex function in $C^2(R^n)$ when A is symmetric and positive definite

CONDITIONS FOR UNCONSTRAINED MINIMUM

- $f: R^n \to R$ and $f \in C^2(R)$

- A necessary condition for the minimum is that at the minimum

  $\nabla_x f(x) = 0$ -> Gradient Vanishes

- A sufficient condition for the minimum is that at the minimum

  $\nabla_x^2 f(x)$ SPD -> Hessian is a (symmetric) positive definite matrix

EQUALITY CONSTRAINED MINIMUM - ELIMINATION

- Method of elimination : Illustration

- Maximize A = ab when 2(a + b) = L is fixed

- Eliminate b in A: b = $\frac{L}{2}$ - a and A = a($\frac{L}{2}$ − a) = $\frac{La}{2}$ − a^2

  $\frac{dA}{da}$ = $\frac{L}{2}$ − 2a and $\frac{d^2A}{da^2}$ = -2 < 0

- At the maximum a*= $\frac{L}{4}$ and b* = $\frac{L}{4}$ and A_{max} = $\frac{L^2}{16}$

EQUALITY CONSTRAINED MINIMIZATION – LAGRANGIAN MULTIPLIER

- Method of Lagrangian multiplier
- Let g: R^n -> R^m and g ∈ $C^2(R^n)$
- $g(x) = (g_1(x), g_2(x), \ldots g_m(x))^T$
- Min f(x) when g(x) = b where b ∈ R^m
- Define the Lagrangian

$$L(x, \lambda) = f(x) + \lambda^T(b - g(x))$$

- λ ∈ R^m is the vector of undetermined Lagrangian multiplier
- At the minimum:

$$\nabla_x L(x, \lambda) = \nabla_x f(x) - \sum_{i=1}^{m} \lambda_i \nabla_x g_i(x) = 0$$

$$\nabla_x L(x, \lambda) = b - g(x) = 0$$

- A necessary condition for the minimum is that at the minimum the gradient $\nabla_x f(x)$ must be a linear combination of the gradients of the constraints

SUFFICIENT CONDITION FOR EQUALITY CONSTRAINTS

- The Hessian of $L(x, \lambda)$ is given by

$$\nabla_x^2 \mathrm{L}(x, \lambda) = \nabla_x^2 f(x) - \sum_{i=1}^{m} \lambda_i \nabla_x^2 g_i(x)$$

- Let $T = \{ y \in R^n \mid y^T \nabla g_i(x) = 0, 1 \leq i \leq m \}$
- T consists of all vectors that are orthogonal to $\nabla g_i(x)$, $1 \leq i \leq m$. Indeed, T is the tangent plane to $g_i(x)$, $1 \leq i \leq m$
- Let x^* be such that there exists $\lambda^* \in R^m$ with
  - *a)* $\nabla_x f(x) = \sum_{i=1}^{m} \lambda_i^* \nabla_x g_i(x)$
  - *b)* $\nabla_x^2 \mathrm{L}(x^*, \lambda)$ is positive definite on T

    $y^T \nabla_x^2 \mathrm{L}(x^*, \lambda) y > 0$ for all $y \in T$

  Then, x^* is a relative constrained minimum

ILLUSTRATION – EQUALITY CONSTRAINT

- Let n = 2, f(x) = $x_1 + x_1 x_2 + \ 3x_2^2$ - to be minimized

  g(x) = $x_1 + 2x_2 \ - 3 = 0$ – constraint
- L(x, λ) = $(x_1 + x_1 x_2 + \ 3x_2^2) - \ \lambda(x_1 {+} 2x_2 - 3)$
- First-order necessary condition:

$$\nabla_x f(x) - \lambda \nabla_x g(x) = \begin{bmatrix} 1 + x_2 - \lambda \\ x_1 + 6x_2 - 2\lambda \end{bmatrix} = 0$$

$$x_1 + 2x_2 \ - 3 = 0$$

- Solution: $x_1^* = 4, x_2^* = -\frac{1}{2}, \lambda^* = \frac{1}{2}$
- $\nabla_x^2 f(x) = \begin{bmatrix} 0 & 1 \\ 1 & 6 \end{bmatrix}$, $\nabla_x^2 g(x) = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$, $\nabla_x^2 L(x, \lambda) = \nabla_x^2 f(x)$
- $\nabla_x f(x^*) = \frac{1}{2}\begin{pmatrix} 1 \\ 2 \end{pmatrix}$, => T = $\{ \frac{2\alpha}{\sqrt{5}} \begin{pmatrix} -2 \\ 1 \end{pmatrix} \mid \alpha \in R\}$
- $\frac{2\alpha}{\sqrt{5}}(-2,1) \begin{bmatrix} 0 & 1 \\ 1 & 6 \end{bmatrix} \frac{2\alpha}{\sqrt{5}} \begin{pmatrix} -2 \\ 1 \end{pmatrix} = \frac{8\alpha^2}{5} > 0$
- Hence ($x_1^* = 4, x_2^* = -\frac{1}{2}$) is a constrained minimum

PENALTY FUNCTION METHOD – EQUALITY CONSTRAINT

- Let f: R^n -> R, g: R^n -> R^m
- Minimize f(x) when g(x) = b
- Consider $P_\alpha(x) = f(x) + \alpha g^T(x)g(x) = f(x) + \frac{\alpha}{2}\sum_{i=1}^{m} g_i^2\,(x)$
- $\alpha > 0$ is called the penalty constant
- $\nabla_x P_\alpha(x) = \nabla_x f(x) + \alpha D_x^T(g)g(x)$

  where $D_x(g) \in R^{m x n}$ is the Jacobian of g(x)
- Let $x^*(\alpha)$ be the solution of $\nabla_x P_\alpha(x) = 0$

PENALTY FUNCTION METHOD – EQUALITY CONSTRAINT

- It can be shown $\lim_{\alpha \to \infty} x^*(\alpha) \to x^*$, the constrained minimum
- Rewrite

$$\nabla_x P_\alpha(x) = \nabla_x f(x) + \sum_{i=1}^{m} \nabla_x g_i(x)[\alpha\, g_i(x)]$$
$$= \nabla_x f(x) + \sum_{i=1}^{m} \nabla_x g_i(x)\lambda_i\,(\alpha)$$

where $\lambda_i(\alpha) = \alpha g_i(x)$, $1 \le i \le m$ plays the role of the Lagrangian multiplier

- It can be shown $\lim_{\alpha \to \infty} \lambda_i(\alpha) = \lambda_i^*$, the value of the Lagrangian multiplier at the minimun

ILLUSTRATION

- n = 2, f(x) = $x_1^2 + \; x_2^2$ - to be minimized

  g(x) = $x_1 + x_2 \; - 1$ - constraint

- $P_\alpha(x) = x_1^2 + \; x_2^2 + \frac{\alpha}{2}[x_1 + x_2 \; - 1]^2$

  $$\nabla_x P_\alpha(x) = \begin{bmatrix} x_1(2+\alpha) + \alpha x_2 - \alpha \\ e x_1 + (2+\alpha) - \alpha \end{bmatrix} = 0$$

  $$\Rightarrow x^*(\alpha) = \left(\frac{1}{2+\alpha^{-1}}, \frac{1}{2+\alpha^{-1}}\right)^T$$

- Multiplier $\lambda(\alpha) = \alpha g(x^*(\alpha)) = \frac{1}{1+\alpha^{-1}}$
- As $\alpha \to \infty$, $x^* = (\frac{1}{2}, \frac{1}{2})^T$ and $\lambda^* = 1$ which is the minimizer obtained using Lagrangian multiplier method

STRONG VS WEAK CONSTRAINED FORMULATION

- Min f(x) when g(x) = b
- Lagrangian multiplier method is called strong constraint formulation which forces the exact equality condition
- Penalty function method is called weak constraint formulation which only forces approximate equality depending on the value of α, the solution is more closer the constraint and $\alpha \to \infty$, the constraint is exactly satisfied
- We will use both of these formulations

EXERCISES

5.1) Plot f(α) = $x(x^2 - 1)$ for $-2 \le x \le 2$ and identify the minima and maxima

5.2) Let f(x) has a minimum at x* then show that af(x), f(x) + c, and af(x) + c all have a minimum at x*

5.3) Find the minimizer of f(x) = $x_1^2 + x_2^2$ when $x_1 + x_2 = 1$ using Lagrangian multiplier method

5.4) Find the x that minimizes $\frac{\alpha}{2}||x||^2$ under the constraint Z = Hx using (a) Lagrangian multiplier and (b) Penalty function method

5.5) Find the minimizer of

1) $f_1(x) = (Z - Hx)^T W (Z - Hx)$

2) $f(x) = f_1(x) + (x - x_b)^T B^{-1} (x - x_b)$

REFERENCES

- D. G. Luenberger (1969) Otimization in Vector Spaces, *Wiley*
- D. G. Luenberger (1973) Introduction to Linear and Nonlinear Programming, *Addison Wesley*
- S. G. Nash and A. Sofer (1996) Linear and Nonlinear Programming, *McGraw Hill*