Neural Networks & Fuzzy Logic (BITS F312)

1. Introduction (Prof. Bhanot)
2. Artificial Neural Networks (Prof. Bhanot)

65%

3. Nontraditional Optimization (Myself)
4. Fuzzy Logic (Myself)

35%

Lecture Plan:

Module1: Nontraditional Optimization (6 Lectures)

Module2: Fuzzy Logic (9 Lectures)

Books:

1. Engineering Optimization by S.S. Rao
2. Evolutionary Optimization Algorithms by Dan Simon
3. Soft Computing by D.K. Pratihar
4. Soft Computing by Jang, Sun and Mizutani
5. Introduction to Type-2 Fuzzy Logic Control by Mendel et al.
6. Fuzzy Control by Hao Ying

Topics to be Covered:

- ✓ Introduction to Traditional Optimization (2 L)
- ✓ Nontraditional Optimization: GA (3 L)
- ✓ Nontraditional Optimization: PSO (1 L)

Introduction to Traditional Optimization

Introduction

- The method to find the 'best' solution out of several feasible solutions

- Express the criterion to judge the goodness of a solution as a function of factors which influence the chosen criterion

- Point at which the function takes its minimum or maximum value is called a minimizer or maximizer respectively

Mathematically,

find $\underline{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$ to (decision/design vector/variables)

minimize/ maximize $f(\underline{x})$ (objective(cost/fitness) function)

subject to

$g_i(\underline{x}) = 0 \quad i = 1,2,\ldots m$ (equality constraints)

$h_j(\underline{x}) \leq 0 \quad j = 1,2,\ldots p$ (inequality constraints)

and

$L_i \leq x_i \leq U_i \quad i = 1,2,\ldots n$ (bounds)

Various Types of Optimization Problems:

1) Unconstrained and Constrained

2) Single and Multiobjective

3) Static and Dynamic

4) Linear Programming and Nonlinear Programming

5) Integer Programming and Real Valued Programming

Traditional & Nontraditional or Metaheuristic Techniques

- Based on pure mathematics
- Yields precise solutions
- May not be suitable to solve complex real-world problems
- They complement each other

- Maximization problems can be converted to minimization problems by modifying the objective function as

$$-f(\underline{x})$$

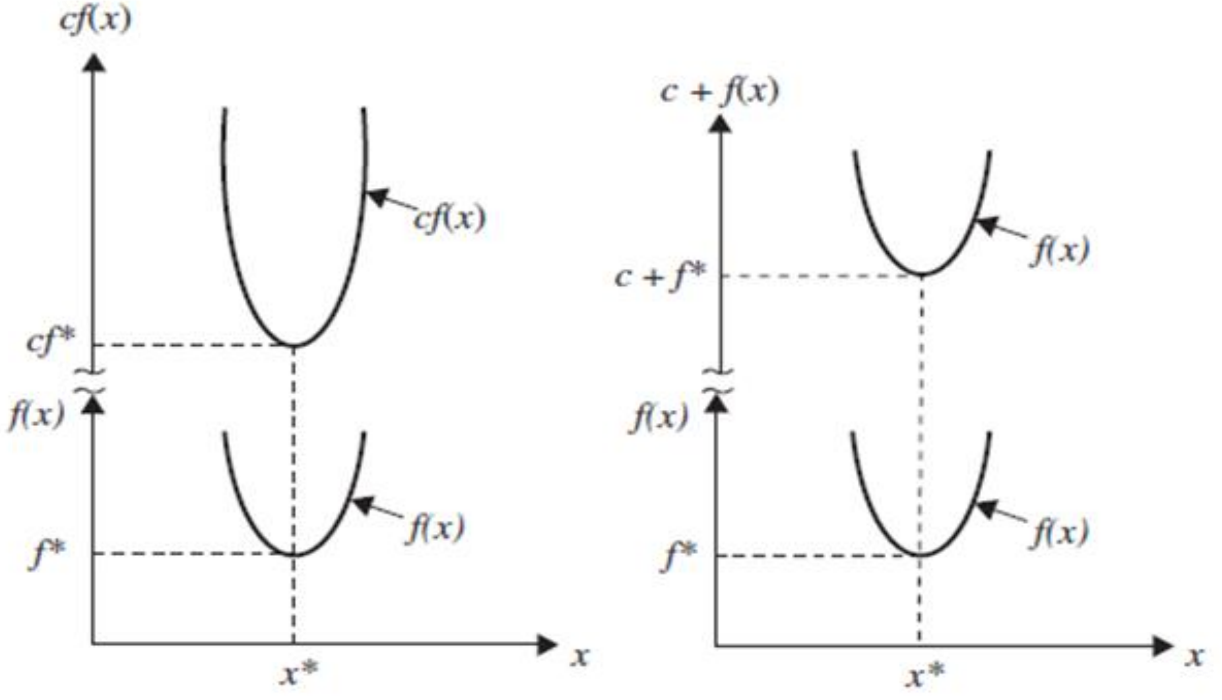
$$\frac{1}{f(\underline{x})} \quad for \;\; f(\underline{x}) \neq 0$$

$$\frac{1}{1+f(\underline{x})} \quad for \;\; f(\underline{x}) \geq 0$$

$$\frac{1}{1+\{f(\underline{x})\}^2} \quad \text{etc.}$$

- Following operations on the objective function will not change the optimum solution $\underline{x}^*$

1. Multiplication/division of $f(\underline{x})$ by a positive constant c.
2. Addition/subtraction of a positive constant c to/from $f(\underline{x})$.



Optimum solution of $cf(x)$ or $c + f(x)$ same as that of $f(x)$

Some Applications

- Optimization problems arise in almost all fields where numerical information is processed (science, engineering, mathematics, economics, commerce, etc.)

- Relevant wherever technical or managerial decision making is involved or a trade off is involved

- Engineering design applications usually have constraints since the variables cannot take arbitrary values and usually we want the design to be best in some sense like minimum cost or maximum efficiency etc.

For example, while designing a bridge, an engineer will be interested in minimizing the cost, while maintaining a certain minimum strength for the structure.

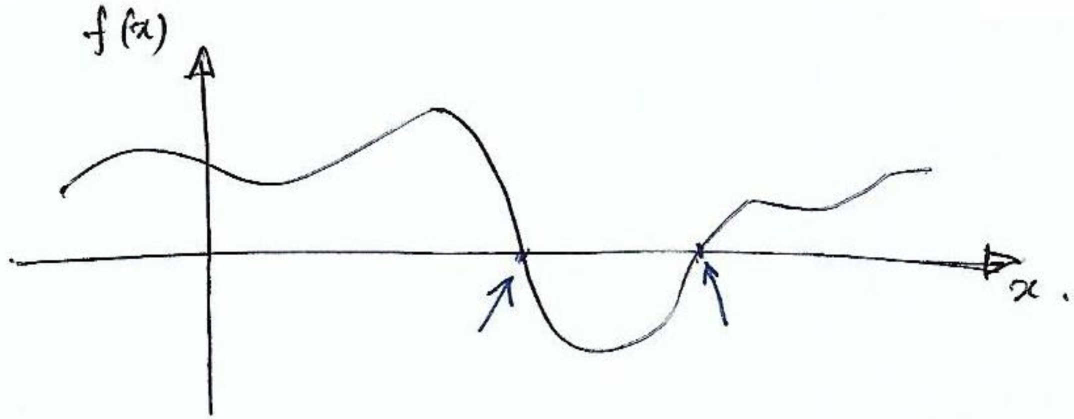
- Control and Signal Processing Applications (E.g. Kalman filter)

- Many other problems can be cast as optimization problems

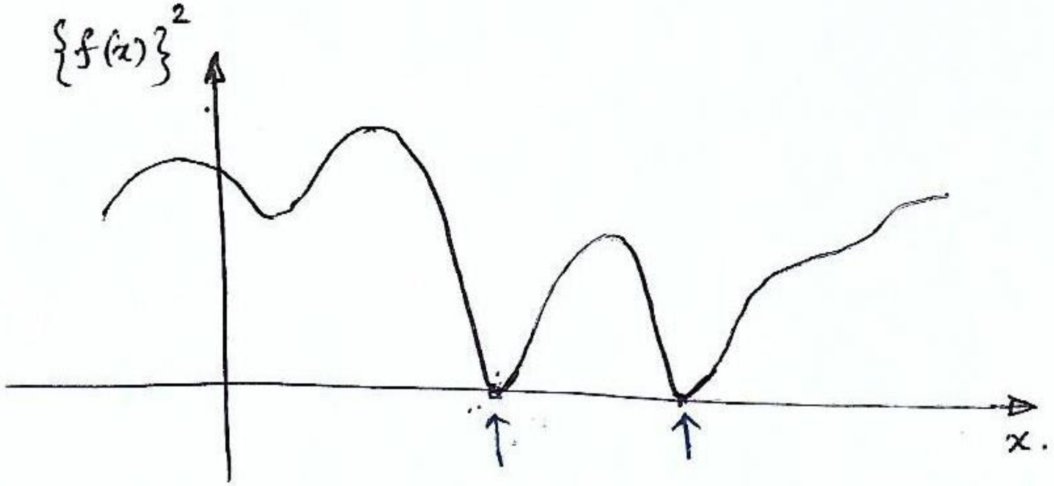
(E.g. root finding; solution of overdetermined/underdetermined linear systems of equations etc.)

Root Finding as an Optimization Problem :

$$f(x) = 0 .$$



minimize $\{f(x)\}^2$



- Over-determined systems
    - More equations and less unknowns

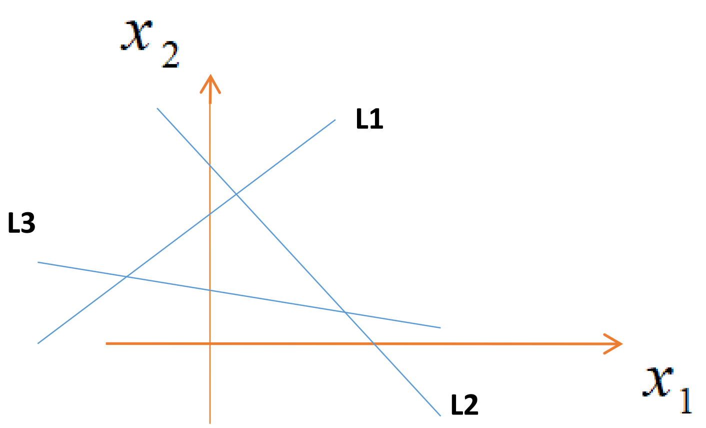
$$a_1x_1 + a_2x_2 + a_3 = 0$$

$$b_1x_1 + b_2x_2 + b_3 = 0$$

$$c_1x_1 + c_2x_2 + c_3 = 0$$

$$A\underline{x} = \underline{b}$$

$$\text{minimize } \|A\underline{x} - b\|.$$

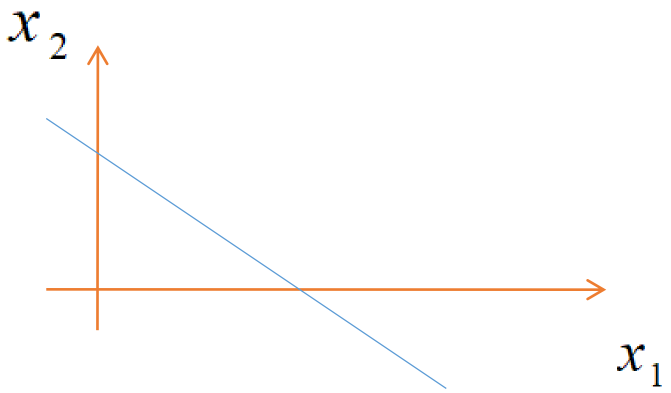


$$\underline{x} = (A^TA)^{-1}A\,.\,\underline{b}$$

- Under-determined systems
  - Less equations and more unknowns

minimize $x_1^2 + x_2^2$

subject to $a_1 x_1 + a_2 x_2 + a_3 = 0$



$$\underline{x} = A^T (A A^T)^{-1} \,.\, \underline{b}$$

A Bit of History

- Originated in Greece during Euclid (~300 BC) (square encloses max. area among all rectangles with the same perimeter)
- Classical analytical methods date back to the era of Cauchy, Bernoulli, Euler and Lagrange
- The beginnings of the subject of **operations research** can be traced to the early period of World War II
- In WW-II, the allied forces faced the problem of allocating limited resources such as fighter airplanes, warships and submarines for deployment to various targets and destinations. Gave birth to linear programming technique.

- During WW-II, Alan Turing used some heuristic search to break the German Enigma codes
- Many developments in the classical techniques during '50s and '60s

- The second wave came in 1970’s partly due to rapid growth in computing power
- Various nontraditional techniques to solve engineering optimization problems started to emerge from 1990’s
- These techniques are mostly inspired from how some biological/natural systems function

Analytical Method: Unconstrained Case

If $f(\underline{x})$ has an extremum at $\underline{x}^*$ then

$$\left.\frac{\partial f}{\partial x_1}\right|_{\underline{x}^*} = \left.\frac{\partial f}{\partial x_2}\right|_{\underline{x}^*} = \cdots \left.\frac{\partial f}{\partial x_n}\right|_{\underline{x}^*} = 0 \quad \text{and}$$

$$\mathbf{H} = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1\,\partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_1\,\partial x_n} \\ \frac{\partial^2 f}{\partial x_2\,\partial x_1} & \frac{\partial^2 f}{\partial x_2^2} & \cdots & \frac{\partial^2 f}{\partial x_2\,\partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_n\,\partial x_1} & \frac{\partial^2 f}{\partial x_n\,\partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_n^2} \end{bmatrix}. \quad = \quad \text{PD or ND} \quad \text{(minimum/maximum)}$$

Saddle point if neither PD nor ND

- **Positive definite matrix(PD):** Determinants of the leading principal minors are all positive

  - for a PD matrix all eigenvalues are positive

$$\begin{bmatrix} a11 & a12 & a13 \\ a21 & a22 & a23 \\ a31 & a32 & a33 \end{bmatrix} \quad \text{is PD}$$

if $a11 > 0$

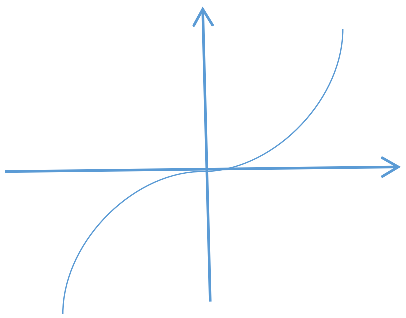
$$\begin{vmatrix} a11 & a12 \\ a21 & a22 \end{vmatrix} > 0$$

$$\begin{vmatrix} a11 & a12 & a13 \\ a21 & a22 & a23 \\ a31 & a32 & a33 \end{vmatrix} > 0$$

- **Negative definite matrix(ND):** Determinants of the leading principal minors are alternately positive and negative with the first one being negative

  - for a ND matrix all eigenvalues are negative

- **Otherwise:** saddle point (E.g. $y = x^3$ at $x=0$)



Analytical Method: Constrained Case (Equality Constraints)

Lagrange Multiplier Method:

minimize $\quad f(\underline{x})$

subject to $\quad g_j(\underline{x}) = 0 \;\; j = 1,2,\dots m$

Lagrangian, $\quad \mathrm{L}(\underline{x}\,,\underline{\lambda}) = f(\underline{x}) + \lambda_1 g_1(\underline{x}) + \lambda_2 g_2(\underline{x}) + \cdots + \lambda_m g_m(\underline{x})$

$$\frac{\partial L}{\partial x_i} = 0 \quad i = 1, 2, \dots\ n$$

$$\frac{\partial L}{\partial \lambda_j} = 0 \quad j = 1, 2, \dots m$$

Physical Meaning of Lagrange Multipliers:

a measure of how rigid the constraints are

$$df^* = \lambda^* \; db$$

Inequality Constraints (Kuhn-Tucker Conditions)

$$g_j(\underline{x}) \leq 0 \qquad j = 1,2, \ldots m$$

$$g_j(\underline{x}) + y_j^2 = 0 \qquad j = 1,2, \ldots m$$

introducing nonnegative slack variables

Checking Second Order Conditions are not as important!