Unconstrained optimization algorithms

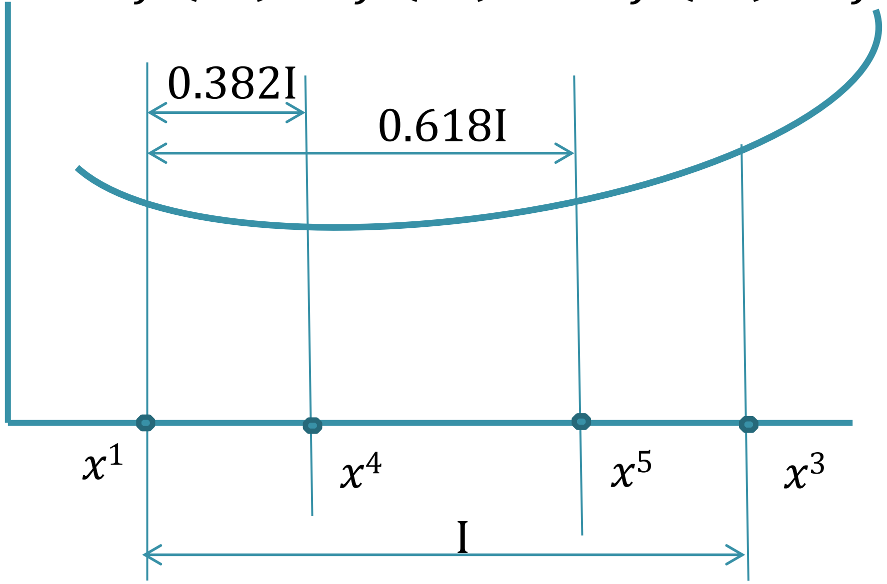
Sridharakumar Narasimhan
Professor,
Department of Chemical Engineering
IIT Madras

Univariate unconstrained optimization

- Minimize f(x)
- **Bracketing:** Find interval containing 3 points such that $f(x^1) > f(x^2)$ and $f(x^2) < f(x^3)$
- Solution lies in interval $x^1 \leq x^* \leq x^3$
- Interval size $\mathrm{I} = x^3 - x^1$
- **Refinement**:
- Golden section method chooses points

$$x^4 = x^1 + 0.382\ \mathrm{I}$$

$$x^5 = x^1 + 0.618\ \mathrm{I}$$



- Evaluate $f(x^4)$ and $f(x^5)$ and identify new interval $[x^1, x^4, x^5]$
- Evaluate at one more point $x^1 + 0.618(x^5 - x^1)$
- Note in either case we need to choose only one more point

Local minimum

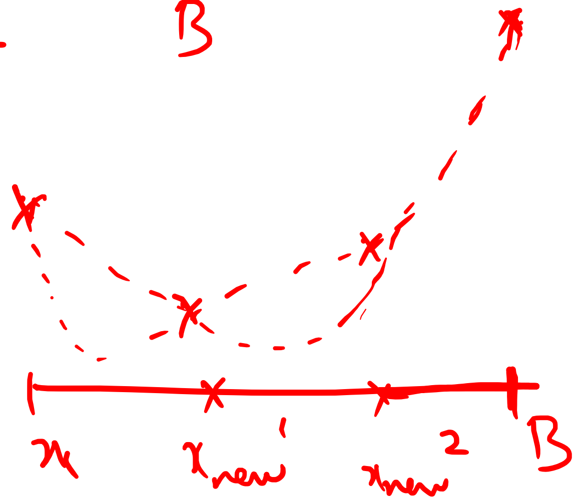
Definitely between x_1 & B.

Eliminate $[A \quad x_1]$

Started with $[A \quad B]$

eliminate $[A \quad x_1]$.

eliminate $[x_{new}^2, \quad B]$

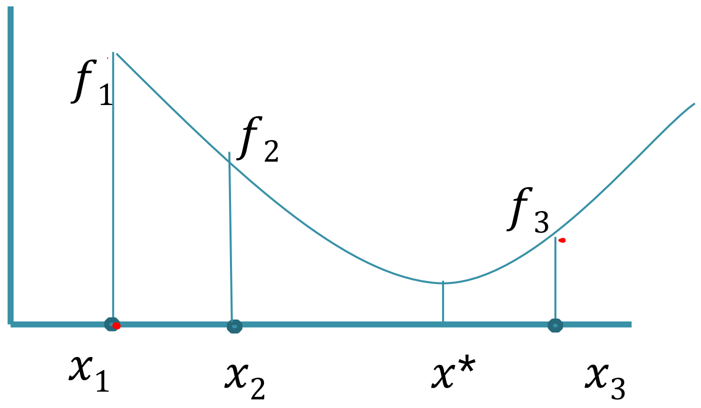


Univariate unconstrained optimization –

- **Refinement based on quadratic or cubic interpolation**
- Fit an approximate quadratic (or cubic) function using three (four) points
  - $f(x) = a + bx + cx^2$
  - Determine constants a, b, c using $[x^1, f(x^1);\ x^2, f(x^2);\ x^3, f(x^3)]$ – will get an exact fit
- Determine optimal solution of quadratic approximation

$$x^* = 0.5\frac{(x_2^2 - x_3^2)f_1 + (x_3^2 - x_1^2)f_2 + (x_1^2 - x_2^2)f_3}{(x_2 - x_3)f_1 + (x_3 - x_1)f_2 + (x_1 - x_2)f_3}$$

- Choose interval consisting of three points such that function decreases and then increase over the interval $[x_2,\ x^*, x_3]$
- Repeat until convergence



Univariate unconstrained optimization - example

- Minimize $f(x) = x^4 - x + 1$
- **Bracketing:**

x	0	-0.5	0.5	1
f(x)	1	1.5625	0.5625	1

- Interval = [0, 0.5, 1]
- **Refinement:**
  - Golden section

x	0	1	0.382	0.618	0.7639	0.5279	
f(x)	1	1	0.6393	0.5279	0.5767	0.5498	

  - Quadratic interpolation

x	0	0.382	1	0.5	0.5728	0.5279	
f(x)	1	0.6393	1	0.5625	0.5348	0.5498	

Multivariate unconstrained optimization

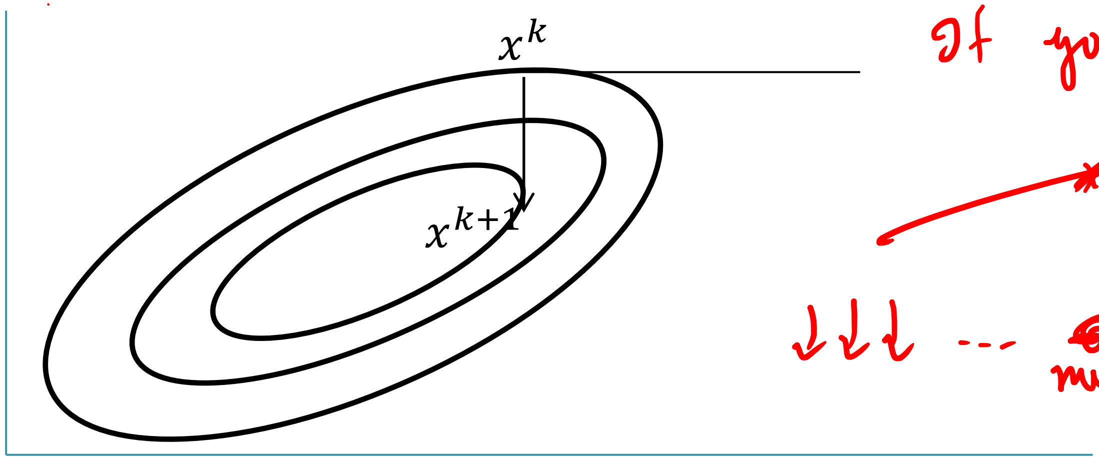
- Minimize $f(x_1, x_2, \ldots x_n) = f(x)$
- **Steepest descent**
  - At iteration k starting point is x^k
  - Search direction s^k = Negative of gradient of f(x) $= -\nabla f(x^k)$
  - New point is $x^{k+1} = x^k + \alpha^k s^k$ where α^k is the value of α for which $f(x^{k+1})$ = f(α) = is a minimum (univariate minimization)

$x^{k+1} = x^k + \alpha(-\nabla f)$

$f(x^{k+1}) = f(x^k + \alpha(-\nabla f))$

$x^k = \begin{bmatrix} x_1^k \\ x_2^k \\ \vdots \\ x_n^k \end{bmatrix}$

$x^{k+1} = \begin{bmatrix} x_1^{k+1} \\ x_2^{k+1} \\ \vdots \\ x_n^{k+1} \end{bmatrix}$



If you go along $-\nabla f$

x^k

$-\nabla f$

min

Steepest descent method - example

- Minimize $f(x_1,x_2) = x_1^4 - 2x_2x_1^2 + x_2^2 + x_1^2 - 2x_1 + 5$
- Initial point $x^0 = [1 \quad 2]$
- $f(x^0) = 5$; $s^0 = -\nabla f(x^0) = [4 \quad -2]$
- $x^1 = [1 \quad 2] + \alpha[4 \quad -2] = [(1 + 4\alpha) \quad (2 - 2\alpha)]$
- Bracketing

0.0796 0.0798

α	0	0.05	0.1	0.0797
x_1	1	1.2	1.6	1.3188
x_2	2	1.9	1.7	1.8406
f(α)	5	4.25	5.1	4.112

$x^1 = \begin{bmatrix} 1 \\ 2 \end{bmatrix} + \alpha \begin{bmatrix} 4 \\ -2 \end{bmatrix}$.

$x_1 = 1 + 4\alpha$.

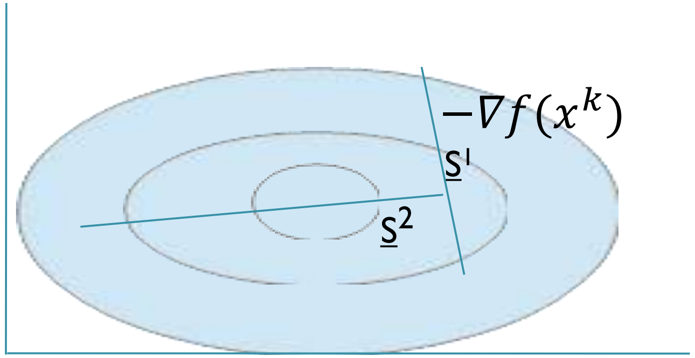
$x_2 = 2 - 2\alpha$.

$f = (1+4\alpha)^4 - 2(1+4\alpha)^2(2-2\alpha) + (2-2\alpha)^2 + (1+4\alpha)^2 - 2(1+4\alpha) + 5$.

Now using bracketing / elimination to identify α:

Multivariate unconstrained optimization

- **Conjugate Gradient Method**
  - s^{k+1} and s^k are H conjugate if $(s^{k+1})^T H s^k = 0$
  - For quadratic function in 2 variables, the optimum is found in 2 steps(iterations) if H is Hessian (matrix of 2nd derivatives) of the objective function



  - Without knowing H:
  - $s^{k+1} = -\nabla f\left(x^{k+1}\right) + s^k \frac{\nabla^T f(x^{k+1}) \nabla f(x^{k+1})}{\nabla^T f(x^k) \nabla f(x^k)}$
  - For a quadratic function in n variables it can be shown that these successive search directions are H conjugate. After n iterations the quadratic function is minimized

Convergence Criteria

- Terminate when $\left\|\nabla f\left(x^k\right)\right\|$ is less than some prescribed tolerance
- Terminate if $||s^k||$ less than tolerance
- Terminate if maximum absolute or relative change in x_i is less than tolerance
- Terminate if relative change in absolute or relative change in objective function f(x) is less than tolerance
- Terminate if maximum number of iterations reached