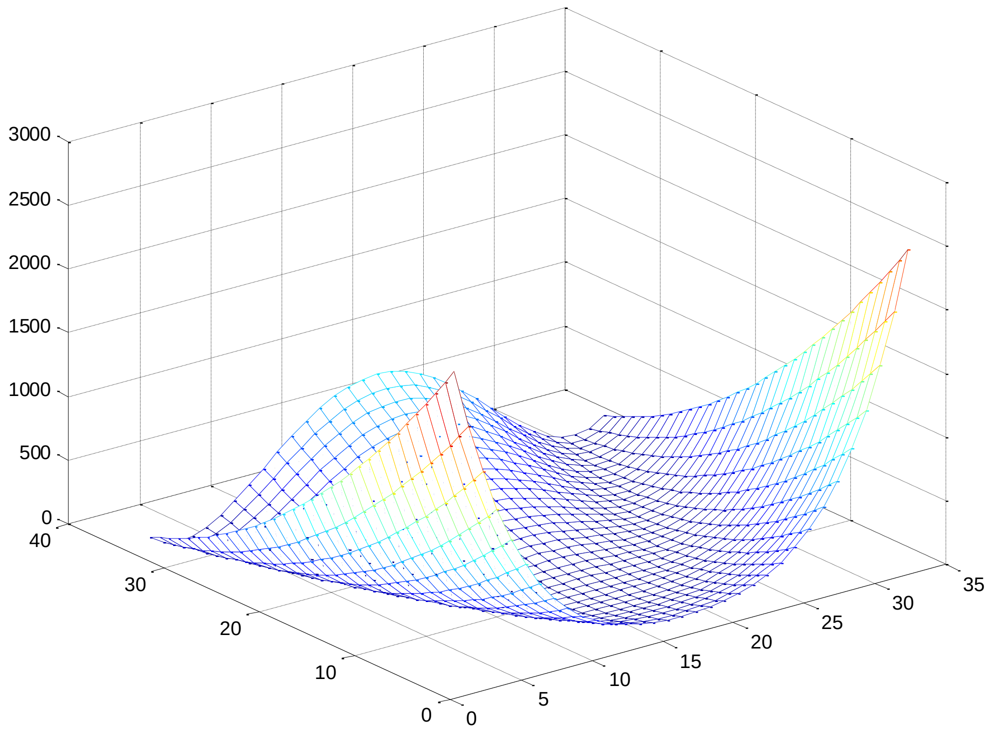
Optimization Methods

A brief discussion on traditional optimization methods

LINKÖPING UNIVERSITY

Contents

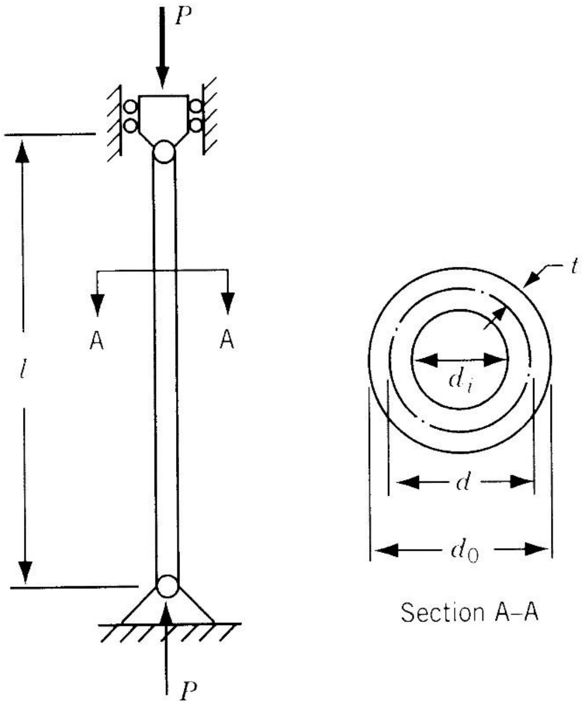
- Graphical solution method
- One dimensional methods
  - Golden section search
  - Fibonacci search
- Gradient based methods
  - Steepest gradient
  - Newton's method
  - Quadratic programming
- MATLAB Optimization toolbox



Graphical solution:
Tubular column design problem

Design a uniform tubular column so that it can withstand the compressive load *P* and simultaneously minimize the cost of the column. Design the column so that the induced stress is less than the yield strength and make sure that buckling does not occur. The cost of the column includes material and manufacturing costs and could be calculated as $5 \cdot m + 200 \cdot d$, where *m* is the mass of the column and *d* is the mean diameter.

P=2500*9.82 N
σ_s=50MPa
E=70000MPa
ρ=2700kg/m^3
l=2.5 m



Graphical solution: Tubular column design problem

The design variables: $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} d \\ t \end{bmatrix}$

The objective function:

$$f(\mathbf{x}) = 5m + 200d = 5\rho\pi dtl + 200d = 5\rho\pi l x_1 x_2 + 200x_1$$

The stress constraint:

$$g_1(\mathbf{x}) = \frac{P}{A} = \frac{P}{\pi dt} = \frac{P}{\pi x_1 x_2} \leq 50 \cdot 10^6$$

Graphical solution:
Tubular column design problem

The buckling constraint: $g_2(\mathbf{x}) = \text{Load} \le \text{Euler buckling load}$

$$g_2(\mathbf{x}) = P \le \frac{\pi^2 EI}{l^2} = \ldots = \frac{\pi^2 E \frac{\pi}{8} x_1 x_2 \left(x_1^2 + x_2^2\right)}{l^2}$$

Design variable limits:

$$0.02 \le x_1 \le 0.14$$

$$0.002 \le x_2 \le 0.008$$

Graphical solution:
Tubular column design problem

min $\quad f(\mathbf{x}) = 5 r p l x_1 x_2 + 200 x_1$

subject to $\quad g_1(\mathbf{x}) = \dfrac{P}{p x_1 x_2 50 \times 10^6} - 1 \pounds 0$

$$\mathrm{g}_2(\mathbf{x}) = 1 - \frac{p^3 E x_1 x_2 \left(x_1^2 + x_2^2\right)}{P 8 l^2} \pounds 0$$

$$\mathrm{g}_3(\mathbf{x}) = - x_1 + 0.02 \pounds 0$$

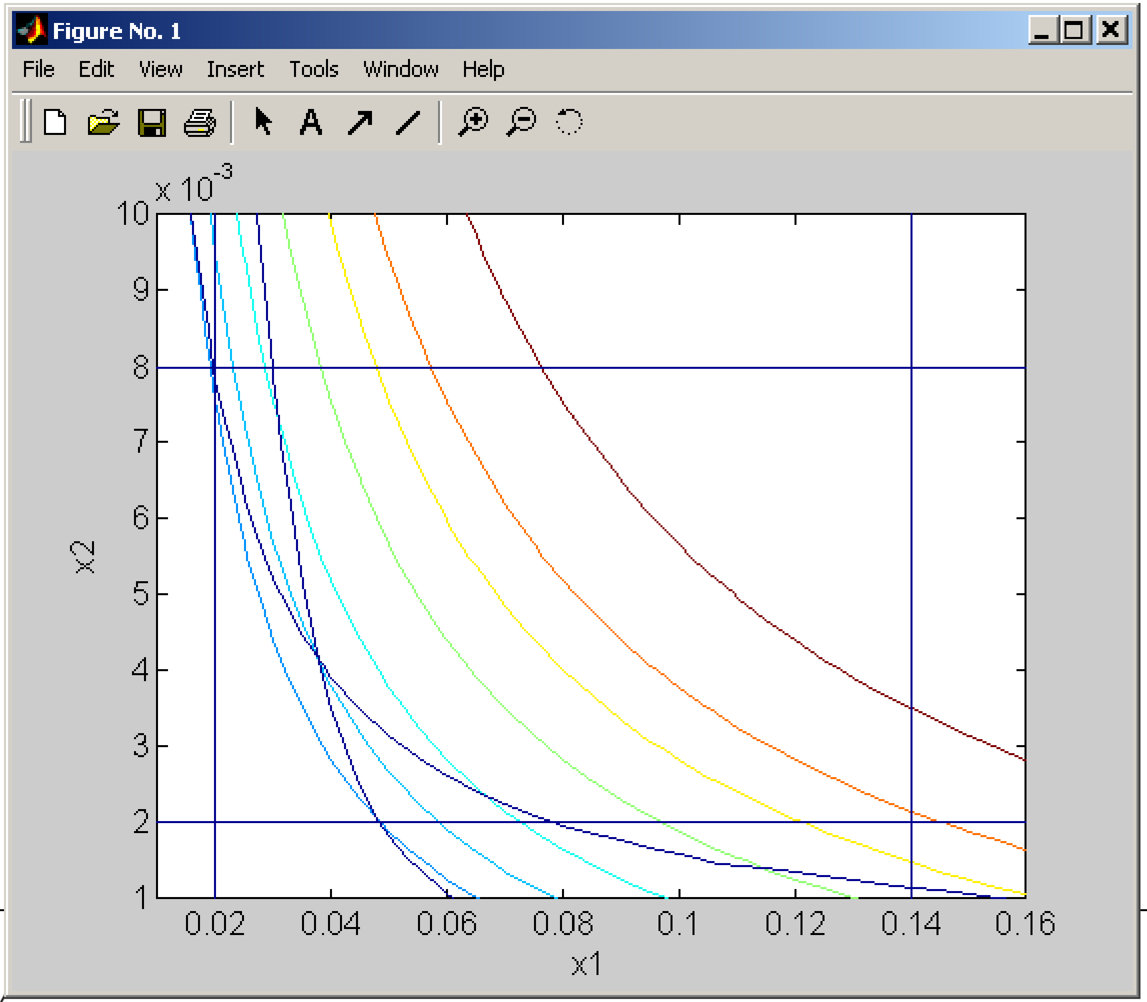
$$\mathrm{g}_4(\mathbf{x}) = x_1 - 0.014 \pounds 0$$

$$\mathrm{g}_5(\mathbf{x}) = - x_2 + 0.002 \pounds 0$$

$$\mathrm{g}_6(\mathbf{x}) = x_2 - 0.008 \pounds 0$$

LINKÖPING
UNIVERSITY

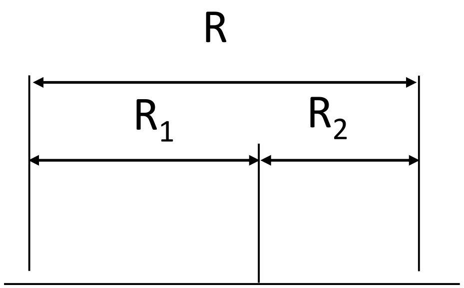
Graphical solution:
Tubular column design problem



LINKÖPING UNIVERSITY

One Dimensional Methods

One dimensional problems: Golden section search



$$\frac{R_2}{R_1} = \frac{R_1}{R}$$

$$R = R_1 + R_2$$

$$\left(\frac{R_2}{R_1}\right)^2 + \frac{R_2}{R_1} = 1$$

$$K = \frac{R_2}{R_1}$$

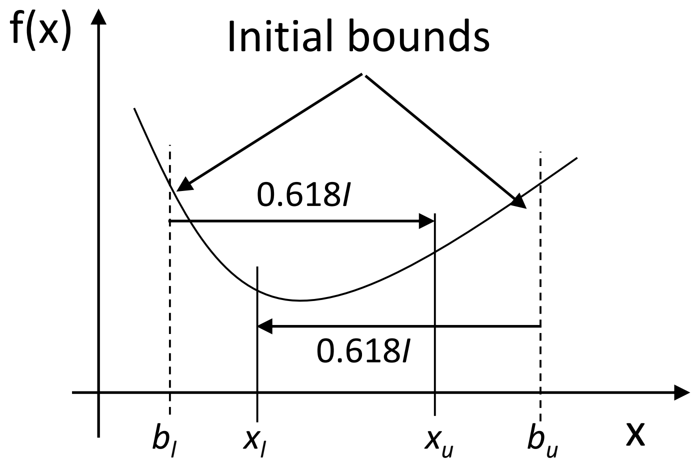
$$K^2 + K = 1$$

$$K = 0.618$$

- Golden section search is an iterative process
- In each iteration the interval of uncertainty is reduced by the value of K=0.618
- After *n* iterations the interval of uncertainty equals:

$$I^n = I_0 \left(0.618\right)^{n-1}$$

Golden section search: procedure



$I = b_u - b_l$ Initial interval

$x_l = b_u - 0.618I$

$x_u = b_l + 0.618I$

The first two points

- Evaluate the function in the two points.
- Identify the point with the lowest value.
- A new interval is created from the boundary and the point with the higher value.
- Create two new points in the new interval of uncertainty.
- The search procedure is now repeated.

Other 1-D methods

- Fibonacci method
  - Similar to golden section
  - Uses Fibonacci numbers as ratios
- Bisection method (intervallhalvering)
  - Uses gradient information
- Approximation methods
  - Fit a polynomial to points and take the derivative of the polynomial.

Approximation Methods – Quadratic Method

The objective function $f(x)$ is approximated with a lower-order polynomial $\theta(x)$.

$$\theta(x) = ax^2 + bx + c$$

Setting $\theta'(x)=0$ yields the minimum value x^* as.

$$x^* = -\frac{b}{2a}$$

LINKÖPING UNIVERSITY

Approximation Methods – Quadratic Method

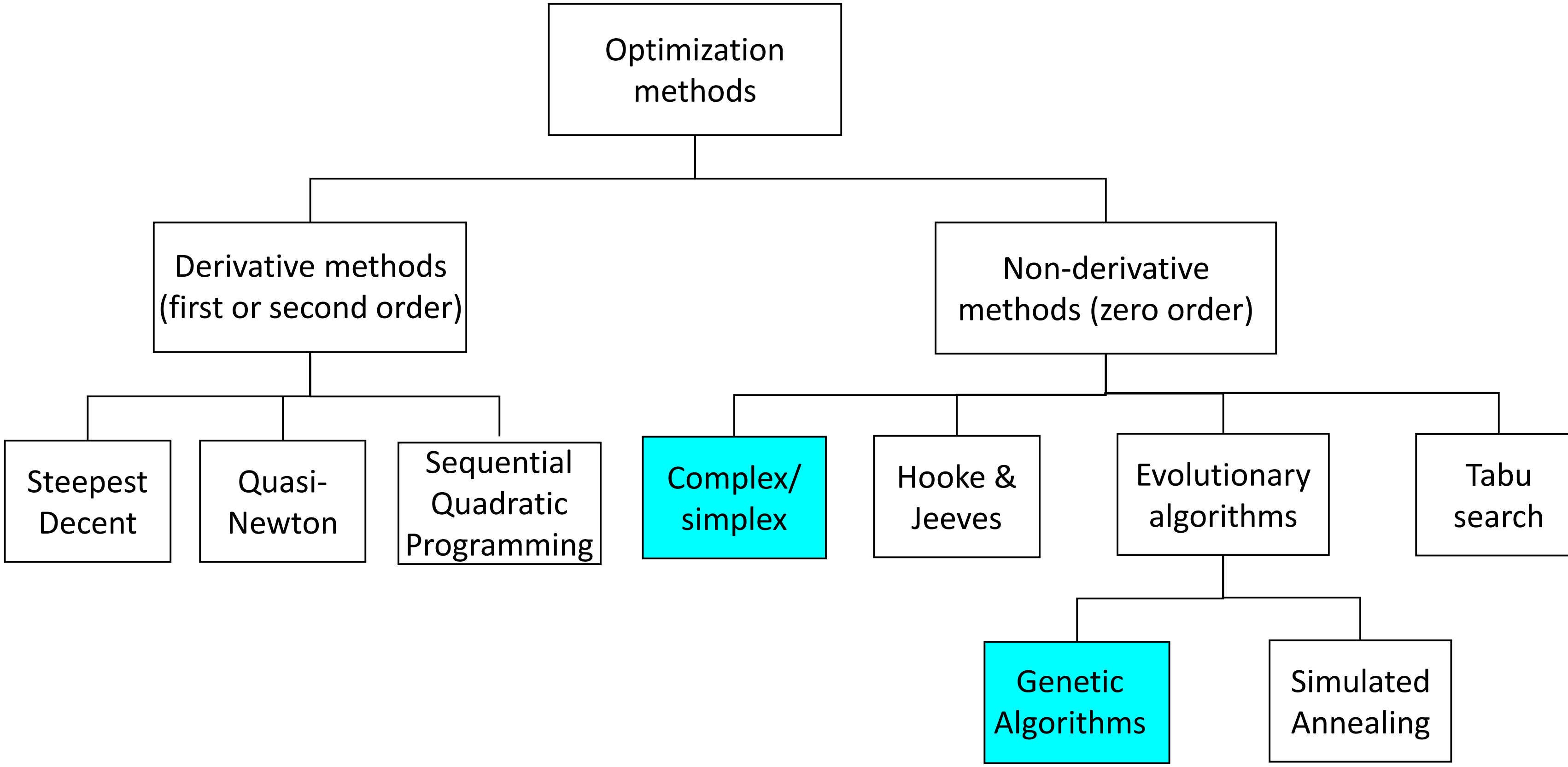
The coefficients *a*, *b* and *c* are calculated by sampling three points (x_1, x_2 and x_3) of the objective function and solving the following equation system.

$$\begin{bmatrix} x_1^2 & x_1 & 1 \\ x_2^2 & x_2 & 1 \\ x_3^2 & x_2 & 1 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} f_1 \\ f_2 \\ f_3 \end{bmatrix}$$

The worst of the three points is then replaced with the new point x^* and new coefficients *a*, *b* and *c* could be re calculated and the process continues.

LINKÖPING UNIVERSITY

Optimization Methods



Derivative Methods

Derivative Methods

- Are classified as first-order or second order depending on whether the first or second order partial derivatives are used.

- Gradient methods are iterative methods involving movement from one point $\mathbf{x}^k$ to another point $\mathbf{x}^{k+1}$ along a given line

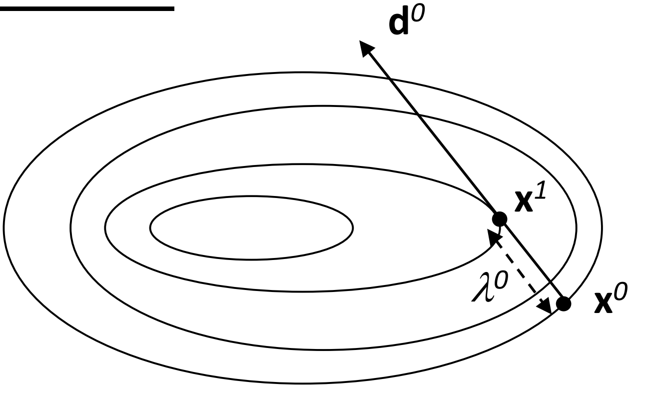
$$\mathbf{x}^{k+1} = \mathbf{x}^k + \lambda^k \mathbf{d}^k$$

$\mathbf{d}$ = *search direction*

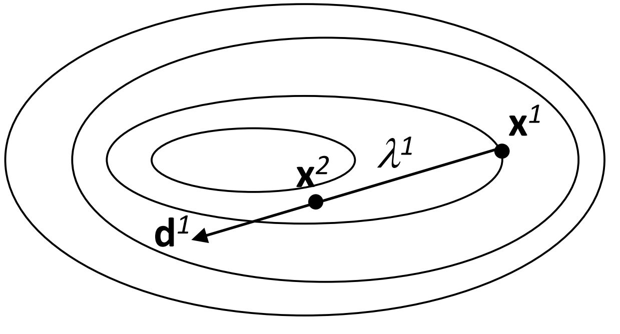
λ = *step size*

Illustration of iterations

First iteration



Second iteration



Steepest descent method

$$\mathbf{x}^{k+1} = \mathbf{x}^{k} + \lambda^{k} \, \nabla f(\mathbf{x}^{k})$$

$\nabla f(\mathbf{x}^{\mathbf{k}})$ = gradient at the point $\mathbf{x}^{k}$

- λ^{k} is determined using a line search to find the minimum of $f(\mathbf{x})$ in $\nabla f(\mathbf{x}^{k})$ direction

- The process is repeated until the step size or the gradient is less then a prescribed tolerance, $\varepsilon>0$.

Steepest descent: procedure

Step 0 Enter starting point x^0 and termination criteria ε. Set $k = 0$.

Step 1 Evaluate the gradient $\nabla f(\mathbf{x}^k)$. If $\left\|\nabla f\left(\mathbf{x}^k\right)\right\| \leq \varepsilon$ then terminate the search and output $\mathbf{x}^k$ as the optimum solution. Otherwise go to step 2.

Step 2 Perform a line search to determine the value of λ that minimizes the function in the direction of the gradient obtained in step 1.

Step 3 Update the search point using $\mathbf{x}^{k+1} = \mathbf{x}^k + \lambda^k(-\nabla f(\mathbf{x}^k))$. Set $k = k+1$ and go to step 1.

LINKÖPING UNIVERSITY

Steepest descent: comments

- Not so popular due to lack of second order derivatives
- Taking the second order derivatives into account improves the performance

Numerical approximation of derivatives

- If the derivatives are not available they could be approximated using finite differences (linear approximation).
- For a 2D problem using forward difference:

$$\left.\frac{\partial f}{\partial x_1}\right|_0 \approx \frac{f\left(x_1^0+\Delta x, x_2^0\right)-f\left(x_1^0, x_2^0\right)}{\Delta x} \qquad \left.\frac{\partial f}{\partial x_2}\right|_0 \approx \frac{f\left(x_1^0, x_2^0+\Delta x\right)-f\left(x_1^0, x_2^0\right)}{\Delta x}$$

- A general central difference is calculated according to:

$$\left.\frac{\partial f}{\partial x_i}\right|_0 \approx \frac{f\left(x_1^0,\ldots,x_i^0+0.5\Delta x_i,\ldots,x_n^0\right)-f\left(x_1^0,\ldots,x_i^0-0.5\Delta x_i,\ldots,x_n^0\right)}{\Delta x_i}$$

Newton's (Second order) method

Consider the Taylor series expansion of the objective function $f(\mathbf{x})$ at the point $\mathbf{x}^k + \Delta\mathbf{x}^k$.

$$f\left(\mathbf{x}+\Delta\mathbf{x}\right)= f\left(\mathbf{x}\right)+\Delta\mathbf{x}^T\nabla f\left(\mathbf{x}\right)+\frac{1}{2}\Delta\mathbf{x}^T\nabla^2 f\left(\mathbf{x}\right)\Delta\mathbf{x}$$

Replace the gradient and the Hessian with **g** and **H** respectively.

$$f\left(\mathbf{x}+\Delta\mathbf{x}\right)= f\left(\mathbf{x}\right)+\Delta\mathbf{x}^T\mathbf{g}\left(\mathbf{x}\right)+\frac{1}{2}\Delta\mathbf{x}^T\mathbf{H}\left(\mathbf{x}\right)\Delta\mathbf{x}$$

Find the $\Delta\mathbf{x}$ that will minimize $f(\mathbf{x})$ by differentiation with respect to $\Delta\mathbf{x}$ and setting the derivative to 0.

$$\Rightarrow \mathbf{g}+\mathbf{H}\Delta\mathbf{x}=0$$

$$\Rightarrow \Delta\mathbf{x}=-\mathbf{H}^{-1}\mathbf{g}$$

Newton’s method

The search equation

$$\mathbf{x}^{k+1} = \mathbf{x}^k - (\mathbf{H}^{-1})^k \mathbf{g}^k$$

$\mathbf{g}^k$ = the gradient at $\mathbf{x}^k$

$\mathbf{H}$ = the Hessian matrix at $\mathbf{x}^k$

In order to effectively cope with objective functions that are not quadratic the search equation is modified and a step size introduced.

$$\mathbf{x}^{k+1} = \mathbf{x}^k - \lambda^k (\mathbf{H}^{-1})^k \mathbf{g}^k$$

Compare to the general search equation

$$\mathbf{x}^{k+1} = \mathbf{x}^k + \lambda^k \mathbf{d}^k$$

Newton's method: procedure

Step 0 Enter starting point $\mathbf{x}^{0}$ and termination criteria ε. Set $k = 0$.

Step 1 Evaluate the gradient $\mathbf{g}^{k}$. If $\mathbf{g}^{k} < \varepsilon$ then terminate the search and output $\mathbf{x}^{k}$ as the optimum solution.

Step 2 Compute Hessian matrix ($\mathbf{H}^{k}$), and its inverse $(\mathbf{H}^{-1})^{k}$.

Step 3 Compute the search direction $d^{k} = (\mathbf{H}^{-1})^{k}\,\mathbf{g}^{k}$.

Step 4 Conduct a line search to determine the λ that minimize $f(\mathbf{x})$ in the search direction $\mathbf{d}^{k}$.

Step 5 Set $\mathbf{x}^{k+1} = \mathbf{x}^{k} + \lambda^{k}\,\mathbf{d}^{k}$. Set $k = k+1$ and return to step 1.

Newton's method: comments

- The method is difficult to implement.
- For complex functions it is hard to compute the Hessian matrix and its inverse.
- Quasi Newton methods use iterative approximation techniques to calculate the Hessian and its inverse.

$$\mathbf{H}^{k+1} = \mathbf{H}^{k} + \hat{\mathbf{H}}^{k}$$

- Davidon – Fletcher – Powel (DFP) and Broyden – Fletcher – Goldfarb – Shanno (BFGS) are examples of such methods.

Quadratic Programming

- Newton's method were developed for unconstrained optimization.

- If the same principle is applied to constrained optimization problems it is called Sequential Quadratic Programming

Matlab Commands

- **linprog** – solves linear programming problems
- **fminbnd** – single variable non-linear bounded minimization using golden section search
- **fminunc** – unconstraint nonlinear minimization
- **fmincon** – constraint nonlinear optimization
- the last two commands uses different methods depending on the problem statement (are gradients and Hessian present).

Questions?